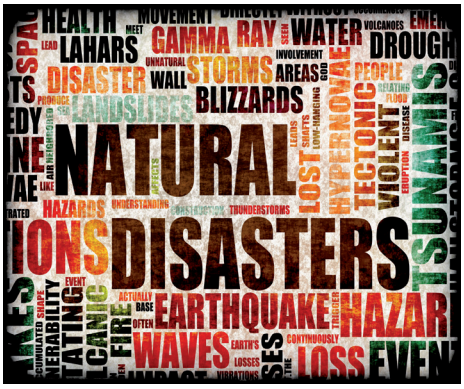
FoKoS
ForschungsKollegSiegen

UNIVERSITÄT SIEGEN

fwu

Mitteilungen des Forschungsinstituts Wasser und Umwelt der Universität Siegen



EVAN 2013

Proceedings of the 1st International Short Conference on Advances in Extreme Value Analysis and Application to Natural Hazards (EVAN 2013)

Heft 6
2013

ISSN 1868-6613

Preface

For more than 30 years I have been working in the field of extreme values, statistical or probabilistic analysis and applications concerning different natural hazards, e.g. storm surges and floods. Developing and applying statistical models for engineering needs is also one of the core research disciplines at the Research Institute for Water and Environment (fwu) of the University of Siegen. The experience we made over the last few years discussing this topic with colleagues from various research fields and stakeholder inspired us to organize and host the first “International Short Conference on Extreme Value Analysis and Application to Natural Hazards (EVAN2013)”. The conference was held in Siegen / Germany from the 16th to 18th September 2013. The importance and current relevance of the conference topic was illustrated only a few weeks before the meeting started by the catastrophic flood along the Danube and Elbe River in the summer of 2013. The conference was jointly organized by the Research Institute for Water and Environment (fwu) and the Institute of Advanced Studies (FoKoS) of the University of Siegen.

Eight outstanding scientists from around the globe were invited to present keynote lectures covering most of the topics covered by the conference. Almost 70 participants from 14 countries attended the conference and listened to nearly 40 presentations.

I want to thank all participants and speakers for their interest and their contribution through interesting discussions to making this a successful event. I also would like to thank the Institute of Advanced Studies (FoKoS) from the University of Siegen, and in particular my colleague Carsten Hefeker, for the financial support to make this conference happen. My thanks also go out to the conference organization committee, namely Jens Bender and his team. They have made a crucial contribution towards the success of the event.

The second “International Short Conference on Extreme Value Analysis and Application to Natural Hazards (EVAN2015)” will be hosted by Fernando Méndez (IH Cantabria) and will take place in Santander / Spain in 2015. I hope to see you all again.

Siegen, October 2013

Jürgen Jensen

Table of contents

Storm Surge Return Periods for the United States Gulf Coast

Hal F. Needham[1], Barry D. Keim[1], David Sathiaraj[1] and Mark Shafer[2]

[1]Louisiana State University, Baton Rouge, LA., USA, Email: hal@srcc.lsu.edu

[2]University of Oklahoma, Norman, OK, USA

Abstract

This study estimates tropical cyclone-generated storm surge levels for 10-year, 25-year, 50-year and 100-year return periods along the U.S. Gulf Coast, using data from SURGEDAT, a global storm surge database. Return periods are calculated using the Pareto, Gumbel and Beta-P distributions, as well as the Huff-Angel and Southern Regional Climate Center (SRCC) linear regression methods. The SRCC method, which fit best with the data, estimated basin-wide surge levels of 8.20 meters for the 100-year return period. The Southeast Louisiana/ Mississippi Zone generated the highest surge levels out of 10 sub-regions in the basin. Surge levels in this zone were 7.67 meters for the 100-year return period. The lowest surge levels in the region were estimated along the West Coast of Florida, where the 100-year storm surge levels are less than four meters. These results are useful for emergency management and disaster science personnel, planners, decision makers and the scientific research community.

1 INTRODUCTION

Storm surges are deadly and costly natural hazards that threaten coastal populations and infrastructure. Despite the deadly and destructive nature of storm surge inundations, limited research has analyzed regional surge statistics, such as the probability of specific surge levels occurring along coastlines. In general, the most common methodologies for such research include storm surge modeling based on historical hurricanes and hurricane frequencies, or surge modeling based on parameterized hurricanes (i.e., Irish *et al.,* 2011). Both of these examples rely on storm surge modeling, which should be verified with empirical data, utilizing surge observations to estimate storm surge quantiles along specific sections of coastline.

Considering the potential threat of surge impacts along the U.S. Gulf Coast, and lack of statistical surge analyses in this region, this hazard is examined through the following objectives: 1) to determine the best method to derive quantile estimates for tropical cyclone-generated surge events along the U.S. Gulf Coast, 2) to calculate Basin-wide surge heights associated tropical cyclones for 2-year, 5-year, 10-year, 25-year, 50-year and 100-year return periods, 3) to parse peak surge data for the Gulf of Mexico into regional clusters, and 4) to calculate the 10-year, 25-year, 50-year and 100-year return periods for the defined regions along the Gulf Coast.

2 METHODS

2.1 Data

This study utilized SURGEDAT, a unique storm surge database developed by Needham and Keim (2012), analyzing the most recent 111 years of SURGEDAT data, from 1900-2010. The 111-year dataset provides 181 storm surges, while only missing 5.7 percent of possible surge events.

2.2 Quantile-estimation methods

Multiple methods were examined to generate quantiles. This study utilized various methods because the performance of quantile-estimation methods varies depending on the characteristics of the

dataset, type of extreme event, and geographic region of the study. Since this study analyzes return periods for storm surge, a research area with little methodological precedent, and these estimates are generated for regions rather than specific locations, we advocate use of linear regression methods. These methods have less restrictive data assumptions than some probability distributions. As such, we selected the Huff-Angel and Southern Regional Climate Center (SRCC) methods as our two primary methods of quantile estimation for these data. The Huff-Angel (Huff & Angel, 1992) and SRCC (Faiers *et al.*, 1997) methods have employed regression techniques to estimate rainfall quantile values in the United States Midwest and South-Central United States, respectively.

In addition, we utilized the Gumbel, Beta-P, and Pareto distributions, primarily for cross-comparison and verification of the linear regression results. The Gumbel distribution was chosen because of its use in Hershfield (1961), a foundational study of heavy rainfall return periods for the United States. Subsequent analysis on heavy precipitation events revealed that the Beta-P method is the best fit for partial duration series of heavy rainfall events in the eastern United States (Wilks, 1993; Wilks & Cember, 1993) and is also utilized in Keim (1998) and Faiers & Keim (2008). The Generalized Pareto distribution introduced by Pickands (1975) was also examined, which is often used to model hydrological extremes over some truncation level (i.e., Hoskin & Wallis, 1987).

Surge levels for associated return periods were provided by fitting a Partial Duration Series (PDS) of surge data to the Huff-Angel and SRCC regression methods of quantile estimation, as well as the Gumbel, Beta-P and Pareto distributions. The Huff-Angel (Huff & Angel, 1992) and SRCC (Faiers *et al.*, 1997) methods both incorporated the Weibull plotting position formula:

$$\text{Exceedence Probability} = \text{Rank} / (n+1) \quad (1)$$

where “n” is the number of years in the data record. Makkonen (2008) provides a detailed discussion of plotting position formulas and their limitations in engineering design. Acknowledging these limitations, the Weibull plotting position formula produced exceedence probability values ranging from .0089 for the surge at Pass Christian, Mississippi during Hurricane Katrina, which was a large-magnitude, less probable event, to .9911 for the low-magnitude, more probable 111^{th}-ranked event. Exceedence probabilities were then utilized to calculate return periods in years, utilizing the formula:

$$\text{Return Period} = 1 / \text{Exceedence Probability}. \quad (2)$$

Using the plotting position formula alone, the surge associated with Hurricane Katrina obtained the longest return period, 112.36 years, while the shortest return period of events in the PDS was only 1.01 years. This means that, on average, we should expect a storm surge of at least 8.47 m to occur somewhere in the Gulf of Mexico every 112 years, while a surge event at least 1.83 m high should be expected annually.

The Huff-Angel linear regression technique, used by Huff and Angel (1992), utilizes a log-log scale (for the x and y axes), thereby graphing and linearizing the PDS surge events in their appropriate Weibull plotting position (Figure 3). The SRCC method, used in Faiers *et al.* (1997), is also a linear regression procedure, utilizing a log scale on the x-axis (return period), and a linear scale on the y-axis (surge height) (Figure 4).
The Gumbel and Beta-P methods also incorporate the same PDS. The Gumbel distribution utilizes the method of moments fitting procedure as in Hershfield (1961), while the Beta-P distribution derives quantile estimates by implementing the Levenberg-Marquardt maximum likelihood fitting procedure (Mielke & Johnson, 1974), as implemented in Wilks (1993).

Surge levels associated with the 2-year return period, utilizing all five methods, ranged from 2.06 meters to 2.86 meters (Figure 5). The Gumbel method calculated the maximum value while the Pareto method calculated the minimum value. The values of the Beta-P, Huff-Angel and SRCC methods fell between these two extremes. Surge levels associated with the 100-year return period ranged from 7.23 meters produced by Gumbel to 13.94 meters estimated by Beta-P (Figure 5). The Beta-P method probably overestimated 100-year return period surge levels, considering the highest actual surge level in the past 131 years was 8.47 meters. The Gumbel method seemingly underestimated the surge levels for the 100-year event, a result that is congruent with Wilks (1993) and Keim & Faiers (2000), who discovered that the Gumbel method commonly underestimates high-magnitude, rare events. The values of the Pareto, Huff-Angel and SRCC methods fell between these two extremes.

The Kolmogorov-Smirnov Statistic (KS-Statistic) was utilized to evaluate the performance of each method. Lower values indicate a better fit to expected results, as noted by Keim & Faiers (2000). The SRCC method produced the best overall results, as it produced the second-best fit for the 111-year analysis and the best fit for the cleaner 82-year analysis. The Pareto method produced the best fit for the 111-year analysis, but the fourth-best fit for the 82-year analysis. Moreover, the Pareto method appears to underestimate surge levels for higher frequency return periods, such as the 2- and 5-year levels, enabling many more surge events to pass this threshold than expected.

3 RESULTS AND APPLICATIONS OF BASIN-WIDE ANALYSIS

The SRCC method estimated return period values ranging from a 2-year surge height of 2.67 meters to a 100-year surge height of 8.20 meters (Table 1). These results will likely benefit various stakeholders with interests along the U.S. Gulf Coast. Basin-wide return period estimates, for example, could help Federal emergency management personnel plan for storm surge disasters, as hazard mitigation and disaster response at that level depend heavily upon risks that are realized at national and regional levels. Basin-wide storm surge return periods also enable scientists to monitor changes in surge frequencies and magnitudes over time, potentially in association with climate change. For example, surge levels could be evaluated against a backdrop of increased future impacts due to sea level rise (Smith *et al.,* 2010).

Table 1: Estimated surge heights associated with return periods for the time period 1900-2010 (111 years) using the Southern Regional Climate Center (SRCC) regression method

Return Period	Surge Level (m)
100-year	8.20
50-year	7.22
25-year	6.24
10-year	4.95
5-year	3.97
2-year	2.67

4 SUB-REGIONAL RETURN PERIOD ANALYSIS

While basin-wide storm surge return periods provide a broad perspective of surge activity in the region, many coastal stakeholders would also benefit from more localized return period estimates. To break the U.S. Gulf Coast into smaller regions, a K-means clustering algorithm was employed. This well-known algorithm minimizes the sum of squared errors (Jain, 2010) which, in this case, was defined as the distance between individual surge observations and cluster centers. In short, K-means assigned each storm surge to a specific cluster, in such a manner that minimized the sum of squared distances between each storm surge and its cluster center.

Given these parameters, the optimal solution divided the Gulf Coast into 10 regions. This procedure provided at least 14 surge observations in each sub-region, while minimizing the size of the sub-regional zones as much as possible. The K-means clustering algorithm assigned each of the 181 surge events to one of 10 clusters, which are named in Table 6 and depicted on a map in Figure 6. The average number of storm surge events in each zone was approximately 18. The Central Texas Zone (Zone 2) contained 14 events, the smallest number, while the Southeast Louisiana/ Mississippi Zone (Zone 5) contained 26 events, the most of any zone.

The Southern Regional Climate Center (SRCC) linear regression method was employed to calculate return periods for these 10 zones because it performed best for the basin-wide analysis. Zones along the Northern Gulf Coast generally observed the highest storm surge levels (Table 2 and Figure 1). The Southeast Louisiana/ Mississippi Zone, which includes the greater New Orleans metropolitan area, observed the highest surge levels for all return periods. The 100-year surge level in this zone is 7.67 meters, while the 10-year return period is 2.72 meters. The Northeast Texas/ Southwest Louisiana Zone, including the Houston-Galveston metropolitan area, observed the

second highest surge levels for each return period, ranging from a 100-year surge of 6.30 meters to a 10-year surge level of 2.44 meters. The lowest surge levels generally occurred along the West Coast of Florida. Surge levels for each return period were lowest in the West Central Florida Zone, which includes the Tampa-St. Petersburg metropolitan area. The 100-year surge level in this zone is only 3.33 meters.

Table 2: Estimated surge heights for 100-yr, 50-yr, 25-yr, and 10-yr return periods in 10 zones along the U.S. Gulf Coast

Zone	10-yr	25-yr	50-yr	100-yr
S TX	1.7	3.21	4.36	5.51
C TX	1.49	3.16	4.42	5.69
E TX/ W LA	2.44	3.98	5.14	6.3
S LA	1.88	2.97	3.8	4.63
SE LA/ MS	2.72	4.69	6.18	7.67
AL/ W FL	2.31	3.8	4.93	6.06
NW FL	1.91	2.56	3.05	3.55
W FL	1.47	2.21	2.77	3.33
SW FL	1.58	2.52	3.23	3.94
FL KEYS	1.74	3.34	4.55	5.77

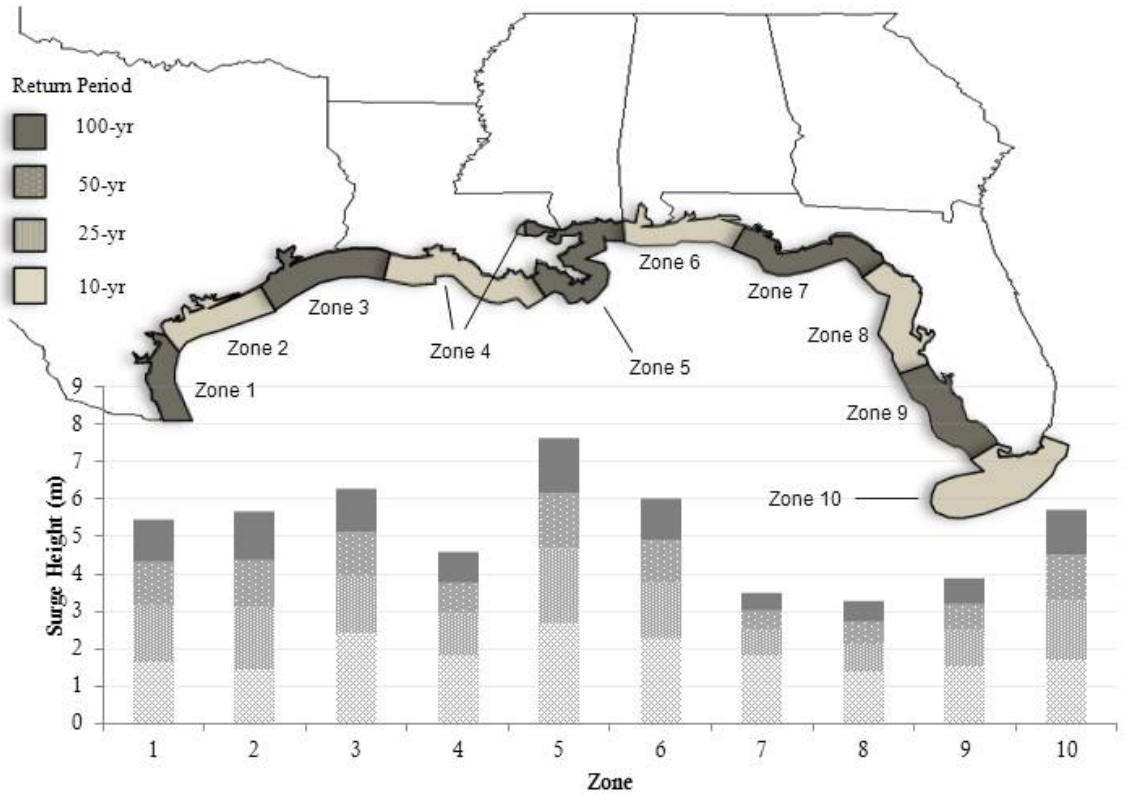


Figure 1: Storm surge levels for 10-yr, 25-yr, 50-yr, and 100-yr time intervals for 10 sub-regions along the United States Gulf Coast

A comparison of storm surge quantiles and hurricane strike frequencies reveals smaller surge return period levels in areas which have observed few strikes from hurricanes and major hurricanes. However, areas that observe frequent major hurricane strikes do not necessarily observe the highest magnitude storm surges. Keim *et al.* (2007) calculated the return period of major hurricane strikes

from 1901-2005 for locations from Texas to Maine, including 19 locations along the Gulf Coast, determining that a portion of Central Florida, from Panama City to Cedar Key, including Apalachicola, is the least likely stretch of Gulf coastline to observe a major hurricane strike. As expected, storm surge quantile estimates for this region are among the lowest along the U.S. Gulf Coast. However, locations most likely to observe strikes from major hurricanes, from Marco Island to Key West, Florida, and secondarily, from Mobile, Alabama to Pensacola, Florida, did not observe the highest storm surge quantiles. Instead, the Southeast Louisiana/ Mississippi zone observed the highest surge quantiles, followed by the Northeast Texas/ Southwest Louisiana zone.

The lack of correlation between frequent major hurricane strikes and enhanced storm surge quantiles makes sense when one considers that many hurricanes impacting South Florida cross the state from east to west, pushing offshore into the Gulf of Mexico. These storms often produce negative storm surges north of the center of circulation, while positive surge levels to the south of the center of circulation generally remain under two meters (Needham & Keim, 2012). Surge magnitudes in these events are relatively small, as the momentum of the storm itself is moving away from the coast, even if winds are blowing onshore to the south of the center of circulation. The majority of hurricanes that impact Texas, Louisiana and Mississippi, however, approach from the south or southeast (Keim & Muller, 2009), providing ample time for these storms to generate high storm surges as they cross the Gulf of Mexico. Therefore, although South Florida observes more major hurricane strikes than locations in the western Gulf of Mexico, storm surge quantiles are generally lower. These observations reveal that the amount of time hurricanes spend over open water before striking a location may be an important factor that influences surge magnitudes.

5 SUMMARY AND CONCLUSIONS

This study conducted a frequency analysis of tropical-cyclone generated storm surge levels along the U.S. Gulf Coast, utilizing 111 years of observed storm surge data. Storm surge quantiles were estimated on a regional level for the entire U.S. Gulf Coast. The Southern Regional Climate Center (SRCC) quantile estimation method produced the best fit results, outperforming the Huff-Angel, Gumbel, Beta-P, and Pareto methods. Basin-wide surge level estimates using the SRCC method ranged from 2.67 meters for the 2-year return period to 8.20 meters for the 100-year return period.

A sub-regional analysis divided the Gulf Coast into 10 segments, enabling the calculation of more localized surge estimates. The K-means clustering method was utilized to parse the Gulf Coast into sub-regions. Quantile estimates were then calculated for these 10 zones using the SRCC method. Three of the four zones with the highest surge levels were located along the northern Gulf Coast, from approximately Galveston, Texas to locations just east of Pensacola, Florida. The Southeast Louisiana/ Mississippi zone has the greatest surge magnitudes, including a 100-year surge level of 7.67 meters. The Northeast Texas/ Southwest Louisiana and the Alabama/ Western Florida Panhandle zones followed with the next highest surge levels; the 100-year surge event in these zones exceeds six meters. In contrast, surge levels were lowest along Florida's West Coast, where the 100-year surge level is estimated at only 3.33 meters.

Sub-regional surge estimates represent the smallest area of analysis in this study. Although coastal stakeholders would benefit from even more localized surge estimates, limited surge data from any given location prevented the possibility of estimating surge quantiles for specific points. Other limitations included underestimating more frequent recurrence intervals, such as the 10- and 25-year surge levels, lack of homogeneity between locations placed in the same region or sub-region, the exclusion of surges generated by frontal systems, particularly in western Florida, and sub-regional clustering according to spatial proximity, which sometimes segregated locations with nearly identical physical geography.

Nevertheless, these results will be useful to a broad audience with interests along the U.S. Gulf Coast. Estimates of extreme storm surge return periods are particularly applicable to emergency management personnel, planners, and decision makers in both public and private sectors. Stakeholders in regional industries, ranging from oil and gas exploration, to port facilities, to seafood production could potentially benefit from this regional assessment of storm surge hazard. The methodology developed in this study could also be utilized to estimate the frequency of storm surge levels in other basins that experience tropical cyclone-generated storm surges.

6 ACKNOWLEDGEMENTS

This research was funded by NOAA Grants NA080AR4320886 and EA133E-07-CN-0084.

7 REFERENCES

Faiers, G.E. and B.D. Keim (2008): *Three-Hour and Twenty-Four-Hour Rainstorm Ratios across the Southern United States.* Journal of Hydrologic Engineering 13 (2), pp. 101-104.

Faiers, G.E., B.D. Keim and R.A. Muller (1997): *Rainfall Frequency/ Magnitude Atlas for the South-Central United States.* SRCC Technical Report 97-1, published by the Southern Regional Climate Center, Department of Geography and Anthropology, Louisiana State University, Baton Rouge, Louisiana, 40p.

Hershfield, D.M. (1961*): Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years*. Technical Paper No. 40, Washington D.C.: National Weather Bureau, 115p.

Hoskin, J.R.M. and J.R. Wallis (1987): *Parameter and Quantile Estimation for the Generalized Pareto Distribution.* Technometrics 29 (3), pp. 339-349.

Huff, F.A. and J.R. Angel (1992): *Rainfall frequency atlas of the Midwest*. Bulletin 71, MCC Research Report 92-03. Published through the Midwestern Climate Center (NOAA), and Illinois State Water Survey, 141p.

Irish, J.L., Resio, D.T. and Divoky, D. (2011): *Statistical properties of hurricane surge along a coast.* Journal of Geophysical Research 116, 15p.

Jain, A.K. (2010): *Data clustering: 50 years beyond K-means.* Pattern Recognition Letters 31 (8), pp. 651-666.

Keim, B., R. Muller and G. Stone (2007): *Spatiotemporal Patterns and Return Periods of Tropical Storm and Hurricane Strikes from Texas to Maine*. Journal of Climate 20, pp. 3498-3509.

Keim, B.D. (1998): *Record Precipitation Totals from the Coastal New England Rainstorm of 20-21 October 1996.* Bulletin of the American Meteorological Society, 79 (6), pp. 1061-1067.

Keim, B.D. and G.E. Faiers (2000): *A comparison of techniques to produce quantile estimates of heavy rainfall in arid and mountainous environments: a test case in western Texas*. Journal of Arid Environments 44, pp. 267-275.

Keim, B.D. and R.A. Muller (2009): *Hurricanes of the Gulf of Mexico*. Louisiana State University Press, 216p.

Makkonen, L. (2008): *Problems in the extreme value analysis.* Structural Safety 30, pp. 405-419.

Mielke, P.W., Jr. and E.S. Johnson (1974): *Some generalized beta distributions of the second kind having desirable application features in hydrology and meteorology*. Water Resources Research 10, pp. 223-226.

Needham, H.F. and Keim, B.D. (2012): *A storm surge database for Gulf of Mexico*. International Journal of Climatology, doi: 10.1002/joc.2425.

Pickands, J. (1975): *Statistical inference using extreme order statistics.* Annals of Statistics 3, pp. 119-131.

Smith, J.M., M.A. Cialone, T.V. Wamsley and T.O. McAlpin (2010): *Potential impact of sea level rise on coastal surges in southeast Louisiana.* Ocean Engineering 37 (1), pp. 37-47.

Wilks, D.S. (1993): *Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series.* Water Resources Research 29 (10), pp. 3543-3549.

Wilks, D.S., and R.P. Cember (1993): A*tlas of Precipitation Extremes for the North-eastern United States and South-eastern Canada*. Publication No. RR 93-5, Ithaca, New York: North-east Regional Climate Center, 40p.

The impact of a 100-year wave and sea level event on a reefed coast

Shari L. Gallop[1,2], Cyprien Bosserelle[2], Ivan D. Haigh[1,2] and Charitha B. Pattiaratchi[2]

[1]Ocean and Earth Science, National Oceanography Centre, University of Southampton, Southampton, United Kingdom, Email: S.Gallop@soton.ac.uk

[2]School of Environmental Systems Engineering, and The UWA Oceans Institute, University of Western Australia, Perth, Australia

Abstract

850 million people live within 100 km of coral reefs, and more than 75 % of the global coastline is rocky. However, there is still relatively little research on the stability and resilience of coasts with rock and coral formations to mean sea level rise and extreme events. As a step to investigate this complex issue, the impact of a 100-year storm was simulated at the reefed coast of Yanchep Lagoon in Western Australia, with present day mean sea level and with an increase of up to 1 m. The beaches of Yanchep are underlain and fronted seaward by calcarenite limestone reefs. The 100-year Annual Recurrence Interval (ARI) levels for waves and tides were defined using numerical hindcasts, validated against observations. The highest water level and wave events measured during the hindcasts were artificially stretched to match the 100-year ARI levels. However, the duration of these events was just 4 days and therefore would not allow enough time for the beach to reach the storm equilibrium profile, so these conditions were run consecutively 3 times. Due to the presence of offshore and nearshore reefs, it was important to ensure that these models included processes such as wave-decay due to bottom friction over reefs. At Yanchep Lagoon, XBeach was modified to significantly decrease model run times and allow increased resolution. This model had the ability to include sand availability maps to allow for rocky outcrops; and can include variable and higher roughness for reefs. The 100-year wave height inside the lagoon was 7.5 % of the offshore wave height. The subaerial beach eroded by up to 3 m, but most of this sand was moved seaward of the reef; or transported alongshore by current jets constricted by the reef to create a sand bank. This sand bank then protected the beach in the lee and limited erosion during the storms. Higher sections of the reef provided increased beach and dune protection.

1 INTRODUCTION

By 2015 half of the world's population will live in coastal zones, including on reefed mainland and island coasts (Hart, 2009). Currently, 850 million people live within 100 km of coral reefs, and more than 275 million live less than 10 km from a reefed-coastline. It is estimated that coral reefs protect shorelines and provide recreational and tourism opportunities in more than 100 countries and territories (Burke *et al.*, 2011). However, globally, more than 75 % of coral reefs are rated as threatened (Burke *et al.*, 2011). In addition to reefs that occur on mainland coasts; they are also integral to the formation and natural maintenance of small island states that are often perceived to be especially vulnerable to extreme events, climate change and mean sea level rise (Woodroffe, 2008) because they are low-lying, have small landmasses and limited freshwater supply; and more than 75% of coral reefs are rated as threatened (Burke *et al.*, 2011).

Current understanding of how reef topography influences spatial and temporal variability in coastal erosion and accretion is limited, although it is clear that rock and reef formations can have a significant impact on the locations and magnitudes of coastal erosion; and the ability of beaches to recover on time-scales from single waves (Bosserelle *et al.*, 2011); through events such as storms (Gallop *et al.*, 2012) and sea breezes (Gallop *et al.*, 2012); to seasonal and inter-annual time-scales (Muñoz-Perez & Medina, 2010). In addition to understanding these processes, to accurately predict the response of reefed coasts to future extreme events, physical and biological influences need to be incorporated. Also, global mean sea levels are expected to rise by about 20 to 80 cm during the

next century (Bindoff *et al.*, 2007). As a step towards understanding and predicting how reefed coasts response to extreme events, the impact of a 100-year storm was numerically simulated on the reefed coast at Yanchep Lagoon in southwestern Australia. To understand the impact of a rise in mean sea level, simulations were also undertaken of the impact of the 100-year storm at present mean sea level and with rises of 0.5 and 1.0 m.

2 STUDY SITE

2.1 Yanchep Lagoon

This research focused on Yanchep Lagoon in southwestern Australia (Figure 1a), where the beaches consist of well-sorted, medium sand (d_{50}=0.4 mm; Murphy, 2011) made of quartz and skeletal material (Semeniuk and Johnson, 1982). Calcarenite limestone reefs outcrop along the coast (Figure 1b) and vary longshore from intertidal and continuous on the southern beaches, to subtidal patches on the northern areas. Southwestern Australia has mainly diurnal tides with a mean spring tidal range of 0.6 m (Department of Defence, 2012). Most storms occur during winter when the wave climate is dominated by swell generated in the Southern ocean. Offshore mean significant wave height (H_s) is 2.14 m and annual mean peak period is 13.5 s and H_s exceeds 4.0 m for 10 % of the year (Bosserelle et al., 2012). A series of three limestone reefs up to 20 km offshore attenuate up to 80 % of the incident swell (Masselink & Pattiaratchi, 2001a).

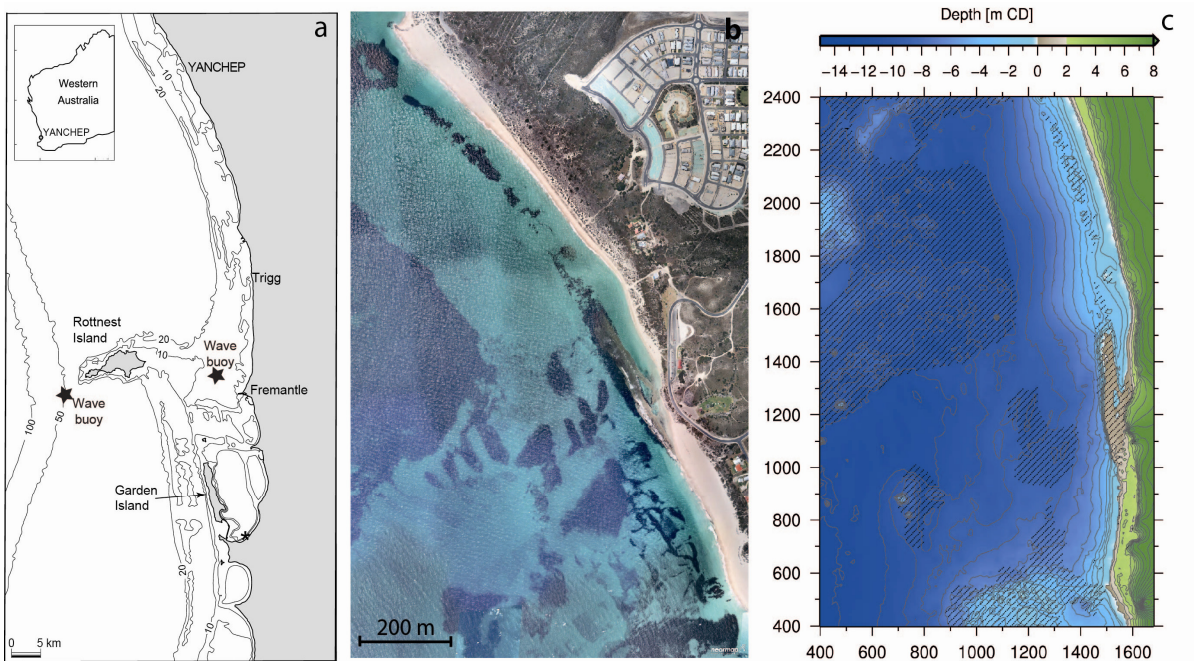


Figure 1: (a) location of Yanchep Lagoon in southwestern Australia; (b) aerial photo of Yanchep showing the limestone reefs (note some dark patches are seagrass wrack) orientated with true north in the vertical (photo source: Landgate); and (c) XBeach model grid orientated with the coast heading north-south showing reefs as hatched areas

3 DEFINING 100-YEAR EXTREME WATER LEVEL AND WAVES

3.1 100-year extreme water level event

Present-day extreme sea level (excluding surface gravity waves) exceedance probabilities around the coast of Australia from combinations of mean sea level, tides and storm surges (extra-tropical and tropical) were estimated from a 61 year-hindcast (1970 to 2009) by Haigh *et al.* (2013a, b). This hindcast utilised a hydrodynamic depth-averaged tide and surge model configured using the Danish Hydraulic Institute's (DHI) Mike21. A flexible mesh was used with resolution of 20 km at the open boundaries down to 2.5 km along the coast and was forced with meteorological fields from the US

National Center for Environmental Prediction's (NCEP) global reanalysis (Kalnay et al., 1996; Kistler *et al.*, 2001). The predicted time-series were validated against measurements from 30 tide gauges and were then fitted to Gumbel extreme value distributions to estimate annual recurrence intervals (ARIs). For the model grid point closest to Yanchep, the 100-year extreme sea level had a height of 1.14 m relative to Australia Height Datum (AHD) where zero is approximately at mean sea level. The highest water level in the time-series of 1.11 m occurred on 16 May 2003 which had a 70 year ARI due to a combination of storm surge occurring at the same time as high tide (Figure 2). In order to use a realistic extreme water level event, surge levels for this event were stretched so that in combination with the tides, water level reached 1.14 m (Figure 2).

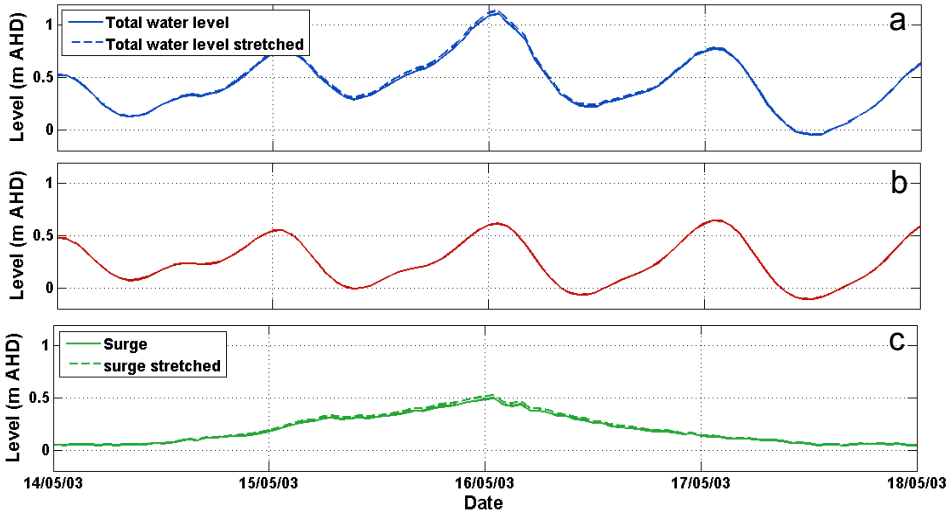


Figure 2: (a) total water level; (b) astronomical tidal water level; and (c) surge level near Yanchep Lagoon. Solid line is the measured levels in May 2003 (70 year ARI) and dashed lines are stretched to have a 100-year ARI

3.2 100-year extreme wave event

Waves in southwestern Australia have only been measured since 1994. Therefore to estimate annual recurrence levels and intervals a 40-year (1970 to 2009) hindcast by Bosserelle et al. (2012) was used, which was validated with measurements from 5 wave buoys along the coast of Western Australia. Wave Watch III (Tolman, 2009) was the numerical modelled used and the model domain had a mosaic of four grids decreasing down to ~15 km resolution along the Western Australian Coast. The model was forced with the NCEP reanalysis wind fields. Similar to the process with water level, at the grid point closest to Yanchep, time-series of annual maxima were fitted with Gumbel extreme value distributions to estimate ARI. The 100-year wave height at Yanchep was estimated as 7.44 m, and the largest event that occurred in the hindcast occurred on 21st July 2009 with a wave height of 7.34 m. Wave height was stretched slightly to peak at 7.44 m (Figure 3).

Incident waves are influenced by shallow-water processes. At Yanchep between offshore and the coast there are three limestone reefs over which some of the wave energy is dissipated. Therefore, the Simulating WAves Nearshore (SWAN) wave model (Booij *et al.*, 1999) was used which includes wave dissipation and breaking. A 10 m-resolution bathymetry was derived from a Light Detection and Ranging Survey (LiDAR) from April 2009. The nearshore wave model grid centred on Yanchep Lagoon, extending 5 km alongshore and 10 km offshore. This model was forced with the 100-year ARI wave event (Figure 3). The offshore limestone ridges dissipated about 75 % of the offshore wave height so that the peak height near Yanchep was 2.04 m (Figure 4).

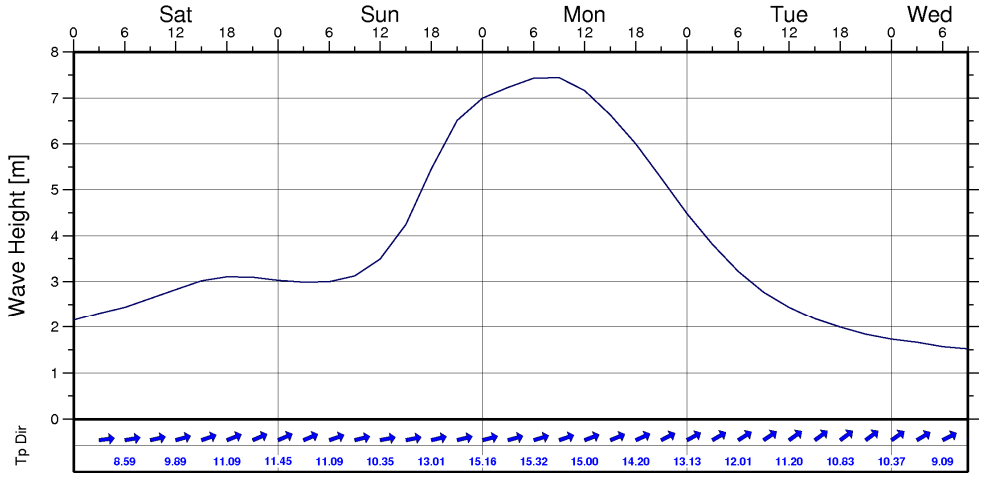


Figure 3: Peak wave period (line) and mean wave direction (arrows) for the 100-year ARI offshore of Yanchep Lagoon

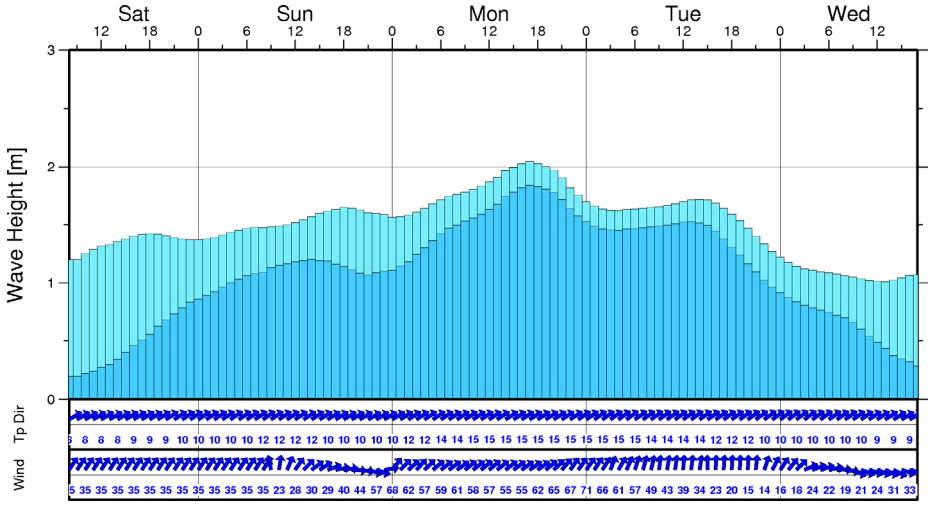


Figure 4: Significant wave height (light blue), swell wave height (darker blue), peak period and mean wind and wave direction for the 100-year ARI event near Yanchep Lagoon

4 IMPACT OF 100-YEAR STORM AT YANCHEP

XBeach (Roelvink *et al.*, 2009) was used to simulate the impact of the 100-year storm at Yanchep Lagoon. XBeach is an open-source numerical model originally developed to model coastal morphology during time-varying extra-tropical and tropical storm conditions. It can include propagation of short-wave envelopes, non-stationary shallow water waves, sediment transport, and bed updating (Roelvink *et al.*, 2010). The model bathymetry and grid had a 5 m resolution, extended about 2 km alongshore and 1 km cross-shore (Figure 1c) and was calibrated using field data collected at Yanchep by Gallop *et al.* (2011, 2012). The calibration process and parameter values used are detailed in Bosserelle *et al.* (in prep). XBeach was forced with the 100-year extreme sea level (Figure 2a) and the 100-year extreme wave event and associated wind (Figure 4). Because of the short duration of the extreme storm, it is likely that the simulated beach would not have had enough time to reach the 100-year storm equilibrium profile. Therefore, the boundary conditions were applied 3 times in a row to simulate the maximum erosion that may occur from a 100-year storm.

At the peak of the 100-year storm, most of the area around Yanchep had currents of less than 0.1 m s-^{1} (Figure 5a). However, in the narrow zone between the reef (Figure 1a and b) and the beach (the coastal lagoon), current speeds reached more than 1 m s^{-1}. The reefs at Yanchep dissipated much of the incident wave height, so that H_s inside the lagoon was less than 0.5 m (Figure 5 b).

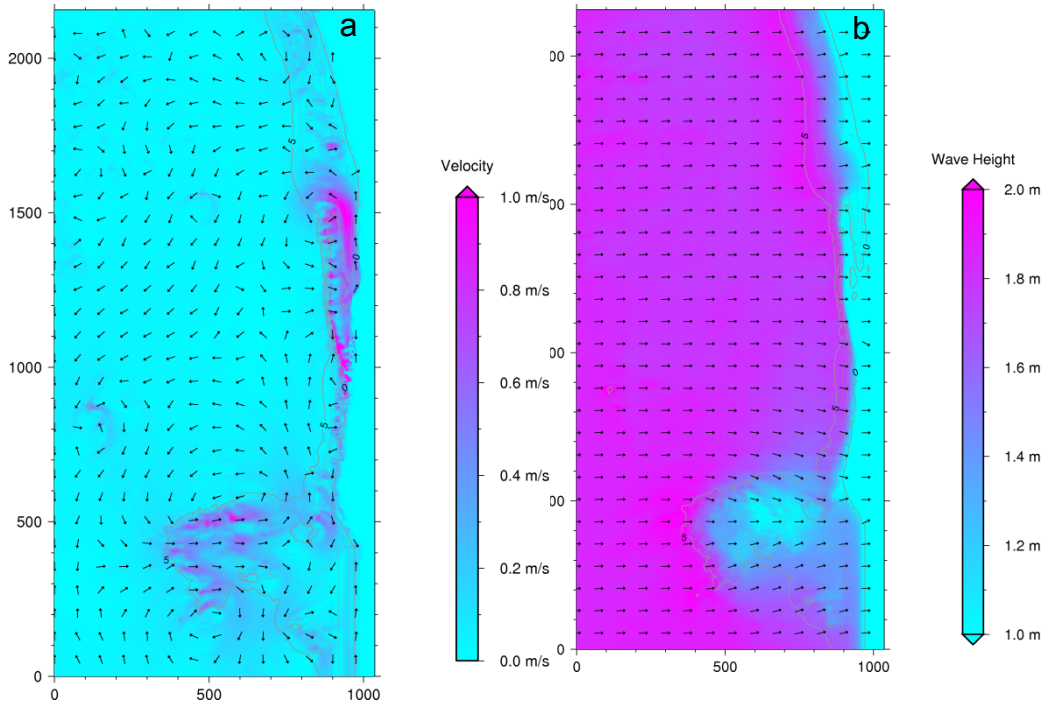


Figure 5: (a) Depth-averaged current speed (colour) and direction (arrow); and (b) wave height (colour) and direction at the peak of the storm

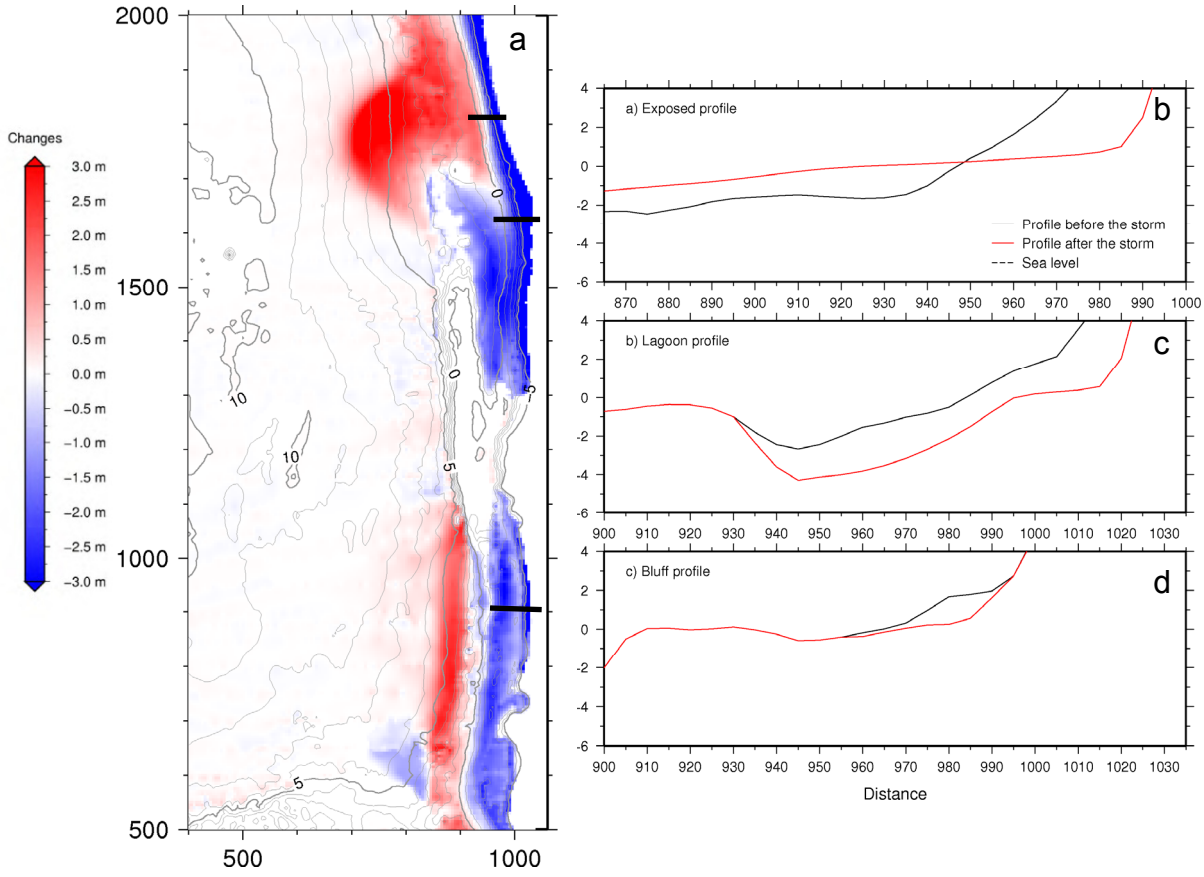


Figure 6: (a) total morphological change after the three 100-year storms; and (b) total morphological change along three profiles shown in (a)

Erosion was predicted to occur along the length of the subaerial beach, except seaward of the limestone headland (Figure 6a). Sand eroded from the southern part of the beach was mostly deposited seaward of the reef. Sand eroded from the middle section of beach and in the lagoon was moved by the current jet (Figure 5a) and deposited as a sand bank north of the single exit of the lagoon. Figure 6b shows that at the exposed area of beach in the north, the dune was eroded by 20 m and sand was deposited offshore. In the lagoonal area, the dune receded by 20 m due to some protection by the reef (Figure 6c). On the southern part of the beach, the dune was stable (Figure 6c). Qualitatively, these results are consistent with measurements made on these same three profiles during a storm with waves with approximately 1 year ARI by Gallop *et al.* (2012).

5 IMPACT OF 100-YEAR STORM AT YANCHEP WITH MEAN SEA LEVEL RISE

To estimate the offshore wave conditions at Yanchep with rises in mean sea level of 0.5 and 1.0 m, the SWAN regional scale wave model was run with the two rises in mean sea level. With higher sea level, there was less interaction with waves and the reefs so that less wave dissipation occurred. Offshore wave heights were predicted to increase by 6 % and 14 % for 0.5 m and 1.0 m rises in mean sea level, respectively (Figure 7a). Inside the coastal lagoon, wave heights were predicted to increase by 18 % and 38 % with the two scenarios of a rise in mean sea level. Higher mean sea level increased erosion of the beach face, with 10 m more horizontal erosion on the beach with both of the mean sea level rise scenarios. This increase in erosion may not be extreme because although with higher water levels these is less wave dissipation, there is also less wave set-up over the reef at Yanchep and a weaker longshore lagoonal current. This means that less sand eroded from the beachface and dunes is exported outside of the lagoon which is likely the main cause of the similar levels of erosion for both scenarios of mean sea level rise.

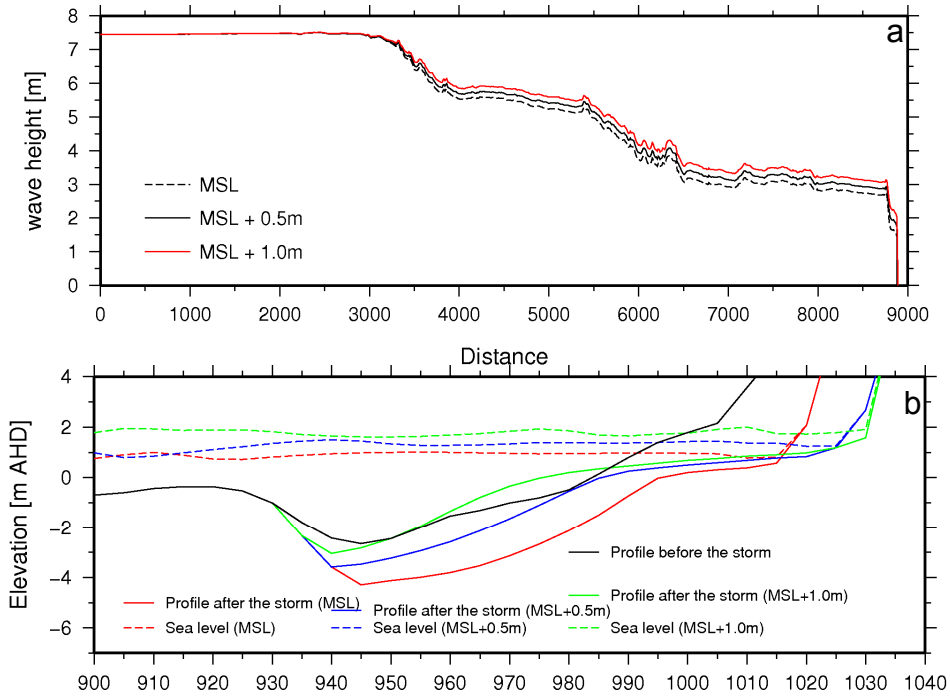


Figure 7: (a) H_s along a cross-shore transect at Yanchep; (b) morphological change during an extreme 100-year storm

6 DISCUSSION: EXTREME EVENTS WITH MEAN SEA LEVEL RISE ON REEFED COASTS

Assessing the impacts of extreme events on any type of coast is challenging. This is particularly true of reefed coasts, and even more so with the addition of mean sea level rise. Reefed coasts come

with many extra challenges and new considerations for predictions of changes in coastal morphology. Some of these challenges are summarised below.

Impact of offshore reefs on nearshore wave conditions

Reefs located up to 10 km offshore of Yanchep dissipated about 75 % of the incoming swell so that the offshore 100-year H_s of 7.44 m was just 2.04 m seaward of these reefs. This shows the importance of an approach that considers not only reefs located along the beach face, but also further offshore during assessment of the impacts of extreme events. In this case, the approach was to use multi-scale numerical modelling, using models appropriate for the scale of interest (e.g. SWAN for the regional scale, and XBeach for the beach scale). A key consideration is that the numerical models take into account wave dissipation such as over the offshore reefs. Without this, there would have been a severe overestimation of the 100-year extreme wave height in the nearshore. Moreover, model bathymetry and grids must be of sufficient resolution to include the effects of reefs on waves.

Current jets

It is generally assumed that reefs protect beaches by dissipating wave energy and result in more stable beaches (e.g. Larson & Kraus, 2000; Vousdoukas *et al.*, 2007). While this may often be the case, predictions of the response of Yanchep Lagoon to a 100-year storm show that this is not necessarily the case on all types of perched beaches during all conditions. One of the key mechanisms for sediment transport at Yanchep is the current jet in the lagoon. This jet transported sediment eroded from the southern part of the beach northwards and deposited it as a sand bank north of the exit of the lagoon. This sand bank means that incident waves in this region would be increasingly more dissipated as the water depth becomes shallower as the sand bank builds up. In contrast, on the southern beach any sediment suspended from the beach face would be transported in the current jet which increased erosion in this area. On coasts such as Yanchep with continuous reefs that extend alongshore, current jets can mean that during extreme events adjacent areas of coast can be erosion 'hot spots' compared to other areas. Generally speaking the > 1 m s^{-1} current predicted in the coastal lagoon (Figure 5a) is of the order of rip current flow speeds (e.g. Austin *et al.*, 2013). However, this lagoonal current jet is usually present at Yanchep, and surface current speeds of up to 1.65 m s^{-1} have been observed during sea breeze winds of 15 m s^{-1} (Gallop *et al.*, 2011). The current speed predicted for the 100-year storm may not be particularly high for this site because although the relatively large waves occurring at the peak of the storm (Figure 5b) could drive faster currents. The high water level during the storm allowed undertow to head seaward out of the lagoon and thus prevented current speeds from getting even higher. Again this shows the need for high-resolution bathymetry and numerical modelling for assessing the impacts.

Additional considerations on coral reefs

The reefs at Yanchep are calcarenite limestone, which is relatively soft. With mean sea level rise, it may be expected that the reefs are impacted even less by waves due to increased water depth. But even so, understanding rates of erosion of rock reefs requires site-specific, detailed measurements. On coasts fronted by reefs made of coral, there are additional considerations when assessing the impact of future extreme events. For example, it is important to include biological effects on the structure of the reef. For example, carbon dioxide concentrations in the atmosphere are expected to exceed 500 parts per million by 2050–2100, and increasing global temperatures by at least 2º (Hoegh-Guldberg *et al.*, 2007). These pressures are expected to lead to ocean acidification which will result in coral bleaching and will reduce carbonate accretion on reefs (Sabine, 2004). This could reduce the size/shape/continuity of the reef structures and therefore decreasing their ability to protect shorelines; and also reduce sediment production which could be catastrophic particularly for small islands where carbonate sediments from reefs are the only sediment source. Quantifying the changes to reef structures and sediment production is likely to be site specific and depend on various factors such as coral species. Therefore, this is a challenge for future research.

7 CONCLUSIONS

Assessing the impact of extreme events, particularly with the effects of mean sea level rise and climate change, is especially challenging on reefed coasts. This research of the impact of a 100-year storm on a reef coasts highlights the need to include effects of offshore as well as nearshore reefs on wave dissipation; and the effect of spatial variability in coastal reef topography on erosion and accretion. The reefs can generate erosion hot-spots, which at Yanchep was largely due to the

formation of longshore current jets constricted by the reefs. On coast fronted by coral reefs, there are also additional challenges to investigating the impact of future extreme events such as including the effect of coral bleaching on reef structures and sediment production.

8 ACKNOWLEDGMENTS

This research was undertaken largely at the University of Western Australia. Apart from the lead author, affiliations of other authors are: C. Bosserelle, Coastal Oceanographer, Secretariat of the Pacific Community (SPC), Applied Geoscience and Technology Division (SOPAC), Suva, Fiji Islands, E-mail: cyprienb@spc.int; I. D. Haigh, Lecturer, Ocean and Earth Science, National Oceanography Centre, University of Southampton, Southampton, United Kingdom, E-mail: I.D.Haigh@soton.ac.uk; C. B. Pattiaratchi, Winthrop Professor, The UWA Oceans Institute, University of Western Australia, Perth, Australia, E-mail: chari.pattiaratchi@uwa.edu.au.

9 REFERENCES

Austin, M.J., Scott, T.M., Russell, P. and Masselink, G. (2013): *Rip current prediction: Development, validation, and evaluation of an operational tool*, Journal of Coastal research 29 (2), pp. 283-300.

Bindoff, N.L., J. Willebrand, V. Artale, A, Cazenave, J. Gregory, S. Gulev, K. Hanawa, C. Le Quéré, S. Levitus, Y. Nojiri, C.K. Shum, L.D. Talley and A. Unnikrishnan (2007): Observations: Oceanic Climate Change and Sea Level. In: *Climate Change 2007: The Physical Science Basis*. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Booij, N., R.C. Ris and Holthuijsen, L.H. (1999): *A third-generation wave model for coastal regions, Part I, Model description and validation*, Journal of Geophysical Research 104, C4, pp. 7649–7666.

Bosserelle, C., Gallop, S.L., Haigh, I.D. and Pattiaratchi, C. (in prep), *Longshore variation in rock topography enhances sand flux on a perched beach in Western Australia*.

Bosserelle, C., Haigh, I.D., Pattiaratchi, C. and Gallop, S. (2011): *Simulation of perched beach accretion using Smoothed Particle Hydrodynamics*, Proceedings of Australasian Coasts and Ports Conference, 28−30 September, Perth, Australia.

Bosserelle, C., Pattiaratchi, C. and Haigh, I.D. (2012): *Inter-annual variability and longer-term changes in the wave climate of Western Australia between 1970 and 2009*, Ocean Dynamics 62, pp. 63−76.

Burke L., Reytar, K., Spalding, M. and Perry, A. (2011): *Reefs at Risk Revisited in the Coral Triangle*, Washington, DC: World Resources 40 Institute, 72p.

Department of Defence (2012): *Australian National Tide Tables 2012*, Australian Hydrographic Publication 11, Canberra.

Gallop, S.L., Bosserelle, C., Pattiaratchi, C. and Eliot, I. (2012): *The influence of coastal limestone landforms on storm erosion and recovery of a perched beach*, Continental Shelf Research 47, pp.16–27.

Gallop, S.L., Bosserelle, C., Pattiaratchi, C. and Eliot, I. (2011): *Rock topography causes spatial variation in the wave, current and beach response to sea breeze activity*, Marine Geology 290, pp. 29−40.

Haigh, I.D., Wijeratne, E.M.S., MacPherson, L.R., Pattiaratchi, C.B., Mason, M.S., Crompton, R.P. and George, S. (in press), *Estimating present day extreme total water level exceedance probabilities around the coastline of Australia: tides, extra-tropical storm surges and mean sea level,* Climate Dynamics.

Haigh, I.D., MacPherson, L.R., Mason, M.S., Wijeratne, E.M.S., Pattiaratchi, C.B. and George, S. (in press), *Estimating present day extreme water level exceedance probabilities around the coastline of Australia: tropical cyclone induced storm surges*, Climate Dynamics.

Hart, D.E. (2009): *The maintenance of reef islands*. Proceedings of the 11th International Coral Reef Symposium, Florida, July 7−11, pp. 409–413.

Hoegh-Guldberg, O., Mumby, P.J., Hooten, A.J., Steneck, R.S., Greenfield, P., Gomez, E., Harvell, C.D., Sale, P.F., Edwards, A.J., Caldeira, K., Knowlton, N., Eakin, C.M., Iglesias-Prieto, R., Muthiga, N., Bradbury, R.H., Dubi, A. and Hatzviolos, M.E. (2007): *Coral reefs under rapid climate change and ocean acidification*, Science 318 (5857), pp. 1737–1742.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, B., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C.,Wang, J., Jenne, R. and Joseph, D (1996), *The NCEP/NCAR 40-year reanalysis projec*t, Bulletin of the American Meteorological Society 77 (3), pp. 437–472.

Kistler, R., Collins, W., Saha, S., White, G., Woollen, J., Kalnay, E., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R. and Fiorino, M. (2001): *The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation*, Bulletin of the American Meteorological Society 82 (2), pp. 247–267.

Larson, M. and Kraus, N. (2000): *Representation of non-erodible (hard) bottoms in beach profile change modelling*, Journal of Coastal Research 16 (1), pp. 1–14.

Masselink, G. and Pattiaratchi, C.B. (2001): *Seasonal changes in beach morphology along the sheltered coast of Perth, Western Australia*, Marine Geology 172, pp. 243–263.

Murphy, P. (2011): *A study of a shoreline salient formed due to a limestone reef at Yanchep Lagoon, Western Australia*, Honours Dissertation, The University of Plymouth, U.K.

Roelvink, D., Reniers, A., Van Dongeren, A., Van Thiel de Vries, J., McCall, R. and Lescinski, J. (2009): *Modelling storm impacts on beaches, dunes and barrier islands*, Coastal Engineering 56, pp. 1133–1152.

Roelvink, D.A., Reniers, A., Van Dongeren, A., Van Thiel de Vries, J., Lescinski J. and McCall, R. (2010): *XBeach Model Description and Manual*, UNESCO-IHE Institute for Water Education, Deltares and Delft University of Technology.

Sabine, C. L. (2004): *The Oceanic Sink for Anthropogenic CO2*, Science 305 (5682), pp.367–371.

Semeniuk, V. and Johnson, D.P. (1982): Recent and Pleistocene beach/dune sequences, Western Australia. Sedimentary Geology 32, pp. 301–328.

Tolman, H.L. (2009): *User manual and system documentation of WAVEWATCH III TM version 3.14*, NOAA/NWS/NCEP/MMAB Technical Note 276.

Vousdoukas, M.I., Velegrakis, A.F. and Plomaritis, T.A. (2007): *Beachrock occurrence, characteristics, formation mechanisms and impact*, Earth Science Reviews 85, pp. 23–46.

Woodroffe C.D. (2008): *Reef-island topography and the vulnerability of atolls to sea-level rise*, Global and Planetary Change 62, pp. 77–96.

Extreme water level statistics for the northern German coastline

Arne Arns[1] and Jürgen Jensen[1]

[1]Research Institute for Water and Environment, University of Siegen, Siegen, Germany, Email: arne.arns@uni-siegen.de

Abstract

One of the main objectives in flood frequency analysis is to estimate heights and occurrence probabilities of extreme events such as floods or storm surges. Currently, different methods are applied on transnational and national scales, resulting in a heterogeneous level of protection. This is why we compare the estimation of extreme still water levels using the main direct methods. From a range of different analyses we provide recommendations that are valid for the German Bight. These recommendations can be used to derive probabilities of extreme still water levels using an automated scheme. However, the availability of water level data is a limitation in many regions. To transfer water level information from gauged to un-gauged sites the concept of regionalization is adopted and adjusted from a riverine to a coastal setting. The regionalization is based on a numerical multi-decadal model hindcast of water levels for the whole of the North Sea. Predicted water levels from the hindcast are bias-corrected using the information from the available tide gauge records. The bias-correction is transferred to the water levels predicted at every coastal and island model grid point in our study area. Using the recommendations on conducting extreme value analyses, water return levels along the entire northern German coastline are estimated.

1 INTRODUCTION

Extreme water levels are a substantial hazard for the low lying countries located in the southern North Sea. In order to plan protection measures against these impacts it is therefore essential to reliably assess extreme sea level. Over the last five decades, several different extreme value analysis methods for estimating probabilities of extreme still water levels have been developed but no universally accepted method has established. Currently, different methods are applied on transnational, but even on national scales, resulting in a heterogeneous level of protection. In Germany e.g., coastal protection is organised by the federal states by defining design water levels using different methods. It is thus difficult to assess the level of protection offered by defences across the different federal states. This is why we compared estimates of extreme water level probabilities using two of the main extreme value analysis methods and conduct a systematic sensitivity assessment of the different steps involved in setting up and using these statistical techniques. The study closes with “best practice” guidance on estimating return water levels in the German Bight.

For this purpose, many stations in the German Bight provide multi-decadal records of high and low waters. However, for some regions (e.g. some islands in the German Wadden Sea) no or just a few water level measurements of any sort exist. As tidal characteristics in the German Bight are highly influenced by shallow water effects and the shape of the coastline, they can differ significantly between stations (see e.g. Jensen & Müller-Navarra, 2008). It is thus difficult to convey information about the likelihood of extreme water level events from gauged to surrounding un-gauged sites. A workaround method is to use the concept of regional frequency analyses (RFA). This concept has widely been applied in hydrology. In RFA it is assumed that flood frequency responses of catchments with similar attributes behave in a similar manner (Merz & Blöschl, 2005). Based on a similarity measure one can decide which information is to be transferred from the catchment to the site of interest. However, this classical approach of regionalization is not applicable in coastal areas (at least not in the German Bight) as water level records show unique characteristics as they are strongly affected by local influences. A regionalization approach therefore needs to account for coastal attributes. This paper investigates the use of a coastal regionalization approach for determining extreme water level probabilities, especially for areas where few or no water level

measurements exist. As a case study we do this for the coastline of Schleswig-Holstein in the German Wadden Sea. In this area, small islands with historical importance are located that could be threatened in the future with sea level rise, thus requiring accurate assessment of flood risk.

2 DATA

Unless stated otherwise, the analyses concerning the best practice of performing extreme value analyses (EVA) in the German Bight (Section 3) are all based on the Cuxhaven record. The Cuxhaven tide gauge is located at the Elbe estuary and provides high and low waters from1843. High resolution data with at least hourly values are available since 1918 (Jensen, 1984; Jensen *et al.*, 1992). In order to investigate the transferability of the results, water level records from 9 further tide gauge stations in the German Bight as given in the EVA column in Table 1 are also considered (see Figure 1 for the location of these). The extension of water level information is conducted using a numerical model of the entire North Sea. A number of tide gauges along the UK, Dutch, French and German coastlines were considered to calibrate (cal.), correct (corr.) and validate (val.) the model generated water levels. To calibrate the numerical model, high resolution tide gauge data along the inner North Sea were used, covering the British East Coast, the English Channel, the Dutch coastline and the German Bight. The calibration was performed using the storm surge event of November 1st, 2006. For the bias-correction of the model output, described below, tidal high water levels for the period from 1970 to 2009 from all German Bight tide gauges (except Pellworm Harbour) were used. Water level records of Pellworm Harbour were used for validation purposes. All water level records were referenced relative to the German ordnance datum “Normalhöhennull” (NHN). The individual records were checked for common errors and suspicious data were deleted.

Table 1: Tide gauges used to evaluate the best practice approach in conducting extreme value analyses (EVA) and to calibrate (cal.), correct (corr.) and validate (val.) the numerical model output. The (*) indicates that tidal high and low waters are available; at all other stations, high resolution values (1-minute) were used. The marker (✓) indicates in which computational step the data was used; stations highlighted with (✓✓) are graphically shown in the paper

Nr.	Tide gauge location (label)	Country	Years considered	Availability [%]	EVA	Model cal.	Model corr.	Model val.
1	Aberdeen (ABE)	GB	2006	100	-	✓✓	-	-
2	Lowestoft (LOW)	GB	2006	100	-	✓✓	-	-
3	Whitby (WHI)	GB	2006	100	-	✓✓	-	-
4	Texel (TEX)	NL	2006	100	-	✓✓	-	-
5	Calais (CAL)	FRA	2006	89.6	-	✓✓	-	-
6	Hörnum (HOR)	GER	2006 1970 – 2009	98.9 100*	- ✓✓	✓✓ -	- ✓✓	 ✓
7	Cuxhaven (CUX)	GER	1918 – 2009	100*	✓✓	✓✓	✓	✓
8	Norderney (HOR)	GER	1935 – 2009	100*	✓	✓✓	✓	✓
9	Helgoland (HEL)	GER	1970 – 2009	100*	-	-	✓	✓
10	Wittdün (WIT)	GER	1970 – 2009	100*	-	-	✓	✓
11	Wyk (WYK)	GER	1970 – 2009	100*	-	-	✓	✓
12	Husum (HUS)	GER	1935 – 2009	99.0*	✓✓	-	✓	✓
13	Dagebüll (DAG)	GER	1935 – 2009	98.9*	✓✓	-	✓	✓✓
14	List (LIS)	GER	1936 – 2009	99.5*	✓	-	✓	✓
15	Büsum (BUS)	GER	1935 – 2009	98.5*	✓	-	✓	✓
16	Schlüttsiel (SCH)	GER	1970 – 2009	100*	-	-	✓	✓
17	Alte Weser (LTA)	GER	1900 – 2009	97.2*	✓✓	-	✓	✓
18	Wilhelmshaven (WIL)	GER	1935 – 2009	99.0*	✓	-	✓	✓
19	Borkum FB (BOR)	GER	1935 – 2009	99.0*	✓	-	✓	✓
20	Emden (EMD)	GER	1970 – 2009	100*	-	-	✓	✓
21	Pellworm Hafen (PEL)	GER	1970 – 2009	100*	-	-	-	✓✓

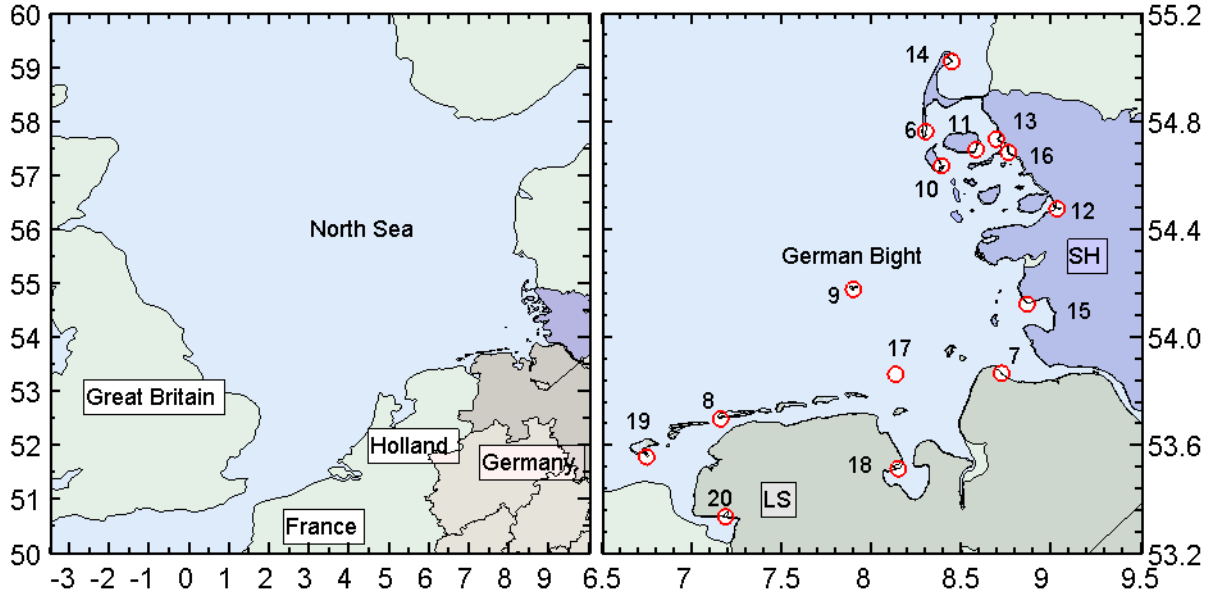


Figure 1: Left: Model domain covering the entire North Sea. Right: Location of tide gauges in the German Bight, tide gauge stations numbering as given in Table 1

3 EXTREME WATER LEVEL ASSESSMENT

3.1 General

The main objectives in flood frequency analysis are to estimate heights and occurrence probabilities of extreme events such as floods or extreme still water levels. In this paper we focus on two direct approaches, namely the block maxima (BM) and the peaks over threshold (POT) methods. The BM method is based on the assumption that the generalized extreme value distribution (GEV) is a good approximation to the distribution of the r-largest water level events within a certain time span (Dixon and Tawn, 1994). The choice of the model determines the way the sample is created. Many studies use the annual maximum (AMAX) value (i.e. r = 1 value/yr), of each year of the record (e.g. Acero *et al.,* 2011). However, it is a wasteful method if further data on extremes are available (Coles, 2001). Further, the 2nd or even the 3rd largest values in a given year can be larger than the AMAX value in another year. Consequently, the AMAX method was extended by Smith and Weissman (1994) (see also Smith, 1986; Tawn, 1988) in order to include a fixed number of independent variables with r > 1 value/yr, the so called r-largest values of each year, into the sample. By applying the r-largest order statistics along the UK coastline, Dixon and Tawn (1994) showed that a choice of r = 8 values/yr appears to yield robust estimates. However, despite incorporating more of the observed extreme data in the estimation of extreme value statistics, even this method can be wasteful if a block does not consider all extremes within a year (Coles, 2001).

The POT approach by contrast is much more efficient (if a not very high threshold is justified) as it considers all values exceeding a certain threshold. Hence, a POT derived sample comprises not only one or a fixed number of events per year. It rather allows for a more rational selection of events fulfilling the criteria of being “extreme” (Lang *et al.*, 1999). In the POT approach, the aim is to develop robust estimates when the model distribution for the exceedances above a threshold is the generalized Pareto distribution GPD (Dupuis, 1998). By comparing AMAX estimates with POT estimates, Cunnane (1973) concluded that the POT approach produces a smaller sampling variance than the AMAX method if the POT derived dataset contains at least 1.65 extreme events per year. The key challenge consists in the determination of an optimal threshold as several important features of frequency modelling are very sensitive to the selected value. If the selected threshold is too low, it causes a bias because the model assumptions are invalid (i.e. values might not be independent or non-extreme data are included in the sample). If the threshold is too high, the variance is large because only few data points are included in the analyses. In extreme value analyses, where models are likely to be extrapolated beyond the observations, this may lead to large

differences in the results. As there is no comprehensive guideline available detailing how to select an appropriate threshold, Arns *et al.* (2013) analysed a broad range of threshold selection methods. They showed that the use of the 99.7th percentile leads to the most stable return water level estimates along the German Bight.

The results of this section are an abstract of the study conducted by Arns *et al.* (2013a); more information on methodologies, data and results are to be found in that study.

3.2 Comparison of methods

The performance of the GEV and the GPD is tested by focusing on the robustness and stability of the particular distribution (see Figure 2). The stability of both methods is evaluated using a water level timeseries that is steadily reduced by one successive year. The last year included is 2008 with the starting year steadily increasing from 1918 to 1998, until the sample reaches a lower limit of 10 years in length. In all cases, the return water level estimates with a return period of 200 years and the associated confidence intervals (CIs) are calculated and plotted against the considered starting time.

To analyse the stability of the BM method, r-largest with r = 1 values/yr are used to create the samples. The stability of the POT method is analysed by applying the 99.7th percentile based threshold, which was identified to be most appropriate for the tide gauges considered here. As shown in Arns *et al.* (2013a), the GEV with r = 1 value/yr is stable when a long record is used. This behaviour changes from1938 onwards, when the GEV derived return water level estimates begin to stagger, with large discrepancies of up to 0.9 m in the resulting return water levels. To obtain reliable and stable return water level estimates for the German Bight using the GEV, we therefore recommend using datasets which start in 1937 or earlier. In this paper, the GEV derived return water levels for the period from 1918 to 2008 are considered as “reference truth”. The stability of the GPD indicates that, in contrast to all cases of the GEV, the GPD leads to very stable return water level estimates until the starting year for the sample creation is in 1977. Using a sample that does not include 1976s values creates unstable results leading to lower return water level estimates. With the starting year in 1997 or later, return water levels increase again.

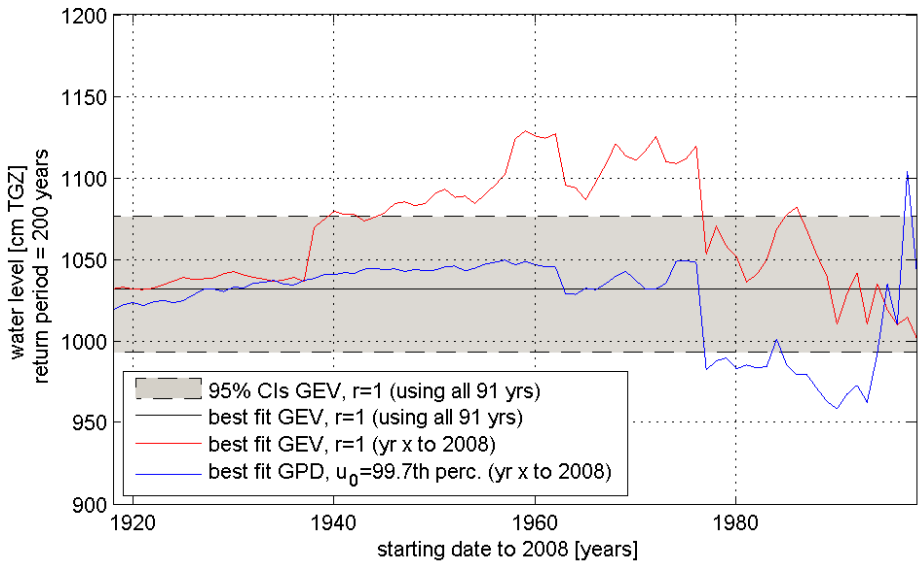


Figure 2: Stability of GPD (U_0=99.7th perc.) and GEV (r=1 val/yr) estimates at Cuxhaven station

In Arns *et al.* (2013a) it was shown that the POT method leads to more stable return water level estimates than the BM method for the Cuxhaven record. To validate this hypothesis for other stations, we analyse a further 9 tide gauge datasets in the German Bight. In the BM method we consider r = 1 value/yr for the GEV and in the POT approach, we use the 99.7th percentile for an automated threshold selection. All return water levels are calculated for a return period of 200 years. Results from applying the GEV and the GPD are shown in Figure 3, for 4 exemplarily chosen sites

(the results are valid at all other stations investigated but not presented here). At all stations, the findings are consistent with the Cuxhaven site, where the GPD performs much more stable than the GEV. As with the Cuxhaven record, all other German datasets show good agreement between the GEV and GPD derived return water level estimates until 1938. Afterwards, the GEV derived return water level estimates begin to fluctuate, causing large discrepancies between the GEV and the GPD. Using a 99.7th percentile derived threshold for analysing the Cuxhaven record yielded negligible differences considering any of the starting years between 1918 and 1976. In the German Bight, this is up to 40-years less than what is recommended for the BM approach (i.e. at least data from 1937 onwards). The findings using the Cuxhaven dataset can thus be confirmed for nine other tide gauge sites in the German Bight.

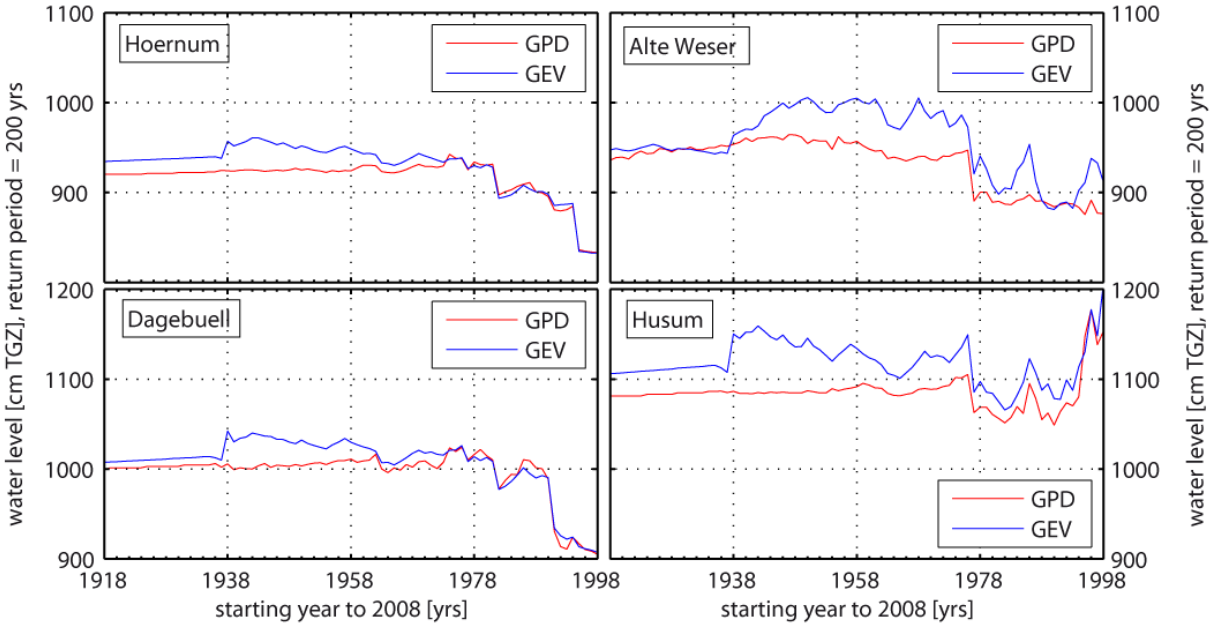


Figure 3: Results of the GPD (U_0=99.7th perc.) and GEV (r=1 val/yr) at the tide gauge records of Hoernum, Alte Weser, Dagebuell and Husum

3.3 Recommendations

An intended outcome of Arns *et al.* (2013a) was to provide guidance for coastal engineers, managers and planners. Based on the sensitivity tests they recommend using a high water peak time series starting in 1976 or earlier as input data and to

- create a stationary dataset using a 1-year moving average trend correction of the high water peaks
- calculate a threshold using the 99.7th percentile of the high water peaks
- use the extremal index for declustering
- fit the GPD to the extreme value sample.

They conclude these recommendations to be valid for the German Bight as well as for the international stations considered in the study (the first point is relevant to just the German Bight), but would benefit from being verified further for other locations around the world.

4 REGIONAL FREQUENCY ANALYSES

The availability of sufficient data is thus one of the crucial aspects when performing statistical analyses. This is why regional frequency analysis (RFA) methods have been developed, enabling to indirectly derive occurrence probabilities. Such methods are based on the assumption that catchments with similar attributes behave similar in flood frequency response and compensate for the lack of data at individual stations (Stedinger *et al.*, 1993) by transferring hydrological information from gauged to somehow related un-gauged target sites. This concept is essentially based on two

steps. The first step is to identify a homogeneous region, suitable to approximately be described by a single distribution that is representative for all sites N considered (Hosking & Wallis, 1993; Rao & Hamed, 2000). In a second step, extreme quantile estimations are conducted. Within homogeneous regions, essential differences between distributions of individual sites are only found in a scaling factor, called the index-flood μ (e.g. mean high water). Exceedance probability Qi(F) at individual sites i=1,…,N may then be calculated by multiplying the regional distribution q(F) with 0 < F < 1 and the index-flood μ according to

$$Q_i(F) = \mu_i q(F) \quad (1)$$

The RFA essentially pursues two objectives. The first one (a) is to enlarge the data basis in gauged areas in order to enhance the precision of flood estimates in the study area. Provided that the considered records are from the same distribution, samples from the joint use of measured at-site data using a number of stations can yield more robust parameter estimates. Using this kind of regionalization represents a substitution between space and time as different long records within an area are used to compensate shorter records (Rao & Hamed, 2000). With respect to practical applications, this concept does not necessarily need to define boundaries between regions but rather includes those sites (or region) being similar to the site (or region) of interest. From mathematical considerations, extreme value samples are considered to describe a random process, comprising independent and identically distributed (IID) random variables. However, water level datasets can exhibit dependencies (so called clusters), which are mostly related to the same meteorological forcing. This dependency has to be taken into account when performing RFA of coastal data sets.

The second objective (b) is to generate information for un-gauged sites. Where information is spatially limited (i.e. little or no data is available in a specific area), regionalization methods can be used to infer hydrologic information from one site (or region) to another. A similarity measure is used to decide which information is to be transferred to the site of interest (Merz & Blöschl, 2005). This task may also be affected by dependencies as the power of statistical homogeneity tests is significantly reduced considering cross-correlated regions (Castellarin *et al.*, 2008).

RFA has often been used in riverine (see e.g. Wiltshire, 1985; add some more) whereas only few studies adjusted this concept to a coastal setting as e.g. for the Dutch coast (van Gelder & Nykov, 1998), the South Chinese Sea (Mai *et al.*, 2006), the French coastline (Bernardara *et al.*, 2011) and the Atlantic and Channel coastline (Bardet *et al.*, 2011). For the German Bight, there is no published study available dealing with a regionalization of water levels. This is why we adjust the use of the most common approach and investigate its application to the German Bight.

With respect to coastal waters it is difficult to define enclosed catchments appropriate for the use in regional flood frequency analyses as distinct boundaries do not exist. This is why we focus on approaches that are based on pooling individual stations into homogeneous groups, where a region describes a set of stations (group) having a similar distribution (Hosking and Wallis, 1993). In literature, several analytical approaches for testing regional homogeneity have been proposed as e.g. the heterogeneity measure H (Hosking & Wallis, 1993; 1997), the discordancy measure Di, (see e.g. Hosking & Wallis, 1991; Rao & Hamed, 2000) or the Wiltshire method (Wiltshire, 1986; for a review and discussion of different methods, see e.g. Viglione *et al.*, 2006; Castellarin *et al.*, 2008). In this study, we use a visual assessment of L-moment dispersion as given in Hosking and Wallis (1993); a method that has been used by other studies that applied RFA to coastal data sets (see van Gelder & Nykov, 1998; Mai *et al.*, 2006; Bardet *et al.*, 2011).

Regional homogeneity is tested using 15 stations covering the entire German Bight. All stations are located in the federal states Schleswig-Holstein (SH) and Lower-Saxony (LS) as given in Figure 1. Occurrence probabilities can be derived by multiplying the regional distribution (GPD) and the index-flood according to Equaton x. In this step, a large degree of uncertainty may be introduced as the index flood from individual stations may have a large variability reflecting the hydrologic diversity within a region (Bocchiola *et al.*, 2003).

The applicability of RFA is tested against at-site analysis by assigning the individual stations to different regions. A first test revealed that combining all stations in the German Bight gives a heterogeneous region. This is why we decided to differentiate at least 2 different regions. From 15

stations in the German Bight (station 6-20 in Table 1) and assuming a region to consist of at least 4 stations we found ~6800 possible homogeneous regions. The entire number of possible combinations amounts to ~15800 (more than half of these combinations yielded heterogeneous regions) as

$$\#\ \text{regions} = \sum_{i=4}^{N} \frac{N!}{k! \cdot (N-k)!} \tag{2}$$

with the total number of regions N and the stations to be drawn without repetitions k. An assessment of the performance of all homogeneous regions was conducted by calculating differences in exceedance probabilities at individual sites from RFA (i.e. multiplying the regional distribution and the index-flood of the individual locations as given in Equation (1)) and at-site analyses. The combination that yielded the lowest root-mean-squared-error (RMSE) was found by assigning stations 6-16 in region 1 (see Table 1) and stations 18-20 in region 2 (see Table 1). This example is shown in Figure 4, where subpanel a) shows differences in region 1 whereas subpanel b) shows differences in region 2.

This comparison highlights differences between quantiles (using the best estimate) to deviate up to 60 cm, with largest discrepancies in higher return periods that reduce with decreasing return periods; an amount quite large with respect to coastal defences. Besides resulting in large discrepancies compared to at-site analyses, applying RFA in coastal areas is subject to another problem. The pooling of regions was done by a comparison of at-site and RFA. There is, however, no clear pattern visibly indicating an objective assignment to regions. With respect to the intended application this leaves the question: which is the region the considered site needs to be assigned to? Up to date we have no similarity measure available helping to overcome this issue. This is why we conclude that it is difficult to convey information about the likelihood of extreme hydrologic events from gauged to un-gauged sites using RFA. Moreover, regionalization approaches (for example by using a numerical model) that account for local storm surge characteristics as e.g. the use of a numerical model are required. Recent applications of extreme value analyses to numerically simulated extreme water level time series can be found in Environment Agency (2011), Haigh *et al.* (2013a, 2013b) and in Arns *et al.* (2013b).

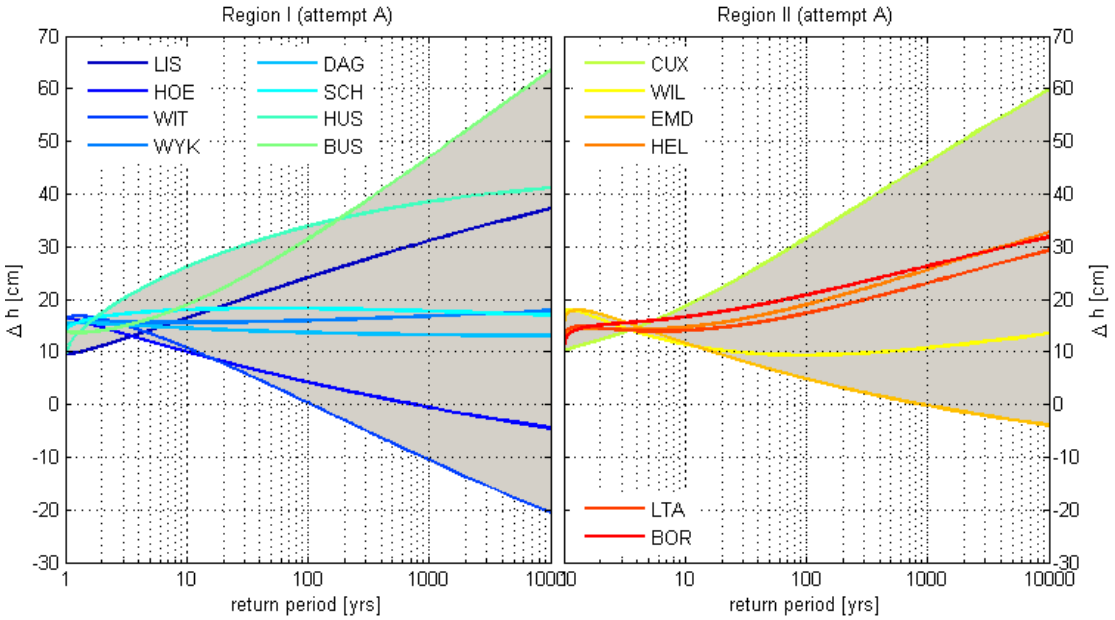


Figure 4: Differences in quantile estimates between RFA and at-site analyses

5 NUMERICAL MODEL

To generate continuous water levels for the entire German Bight, a hindcast for the 40-year period from 1970 to 2009 was performed with a hydrodynamic-numerical model. A two-dimensional, depth-averaged barotropic tide-surge model of the entire North Sea has been configured using the Danish Hydraulic Institute's (DHI) Mike21 FM (flexible mesh) model suite. The software is based on the numerical solution of the incompressible Reynolds averaged Navier-Stokes equations; the spatial discretization is achieved using a flexible mesh. The model was configured within a coastline provided by the National Oceanic and Atmospheric Administration (NOAA) with a resolution of 1:250.000 (http://www.ngdc.noaa.gov/mgg_coastline/). The resolution of the coastline was resampled to 30 km along the open boundaries, increasing to 10 km in the northern- and southernmost parts of the European mainland coastline. In between these locations (Scandinavia, the Netherlands, Belgium, France), the resolution was successively resampled until reaching a maximum resolution of 1 km in the German Bight. The bathymetric data, interpolated onto the model grid, was obtained from a range of different sources including high resolution (~ 15 m) survey maps of the Wadden area provided by the Schleswig-Holstein's Government-Owned Company for Coastal Protection, National Parks and Ocean Protection (LKN-SH), a bathymetric chart with a resolution of 1 nautical mile provided by the Federal Maritime and Hydrographic Agency (BSH) and the General Bathymetric Chart of the Oceans (GEBCO) data provided by the British Oceanographic Data Centre (BODC) with a resolution of 0.5°. All datasets were corrected to the German reference datum NHN.

At the open boundaries, the model was driven by astronomical tidal levels. These were derived from a global tide model provided by MIKE21 (DHI), including the eight primary harmonic constituents (K1, O1, P1, Q1, M2, S2, N2 und K2, see e.g. Andersen, 1995). Additionally, the Mean Sea Level (MSL) was considered using an index-time series for the entire North Sea from Wahl *et al.* (2013); the time series was derived using data from 30 tide gauges located around the North Sea basin. As each year of the considered 40-yr hindcast was run separately, the MSL at the open boundaries was adjusted according to the annual MSL time series. The surge component of the model was generated by forcing the model with mean sea level pressure fields and u and v components of 10 m wind fields provided by the CIRES 20th Century Reanalysis V2 Project (Compo *et al.*, 2011) of the Earth System Research Laboratory, US National Oceanic & Atmospheric Administration (NOAA). These fields are available with a spatial resolution of 2° and a temporal resolution of 3 hours. The bed resistance was set to a constant Manning's number of $n = 0.022$ [-] (corresponds to $k_{st} = 45\ m^{1/3}/s$). The model was run for a two day warm up period and results were stored with at an interval of 10 minutes. The model was calibrated using stepwise variations of the bed resistance, using Manning's n-values between $0.022 \leq n \leq 0.028$ [-].

5.1 Bias-correction

The calibration exercise allowed us to minimise the differences between the observed and the modelled water levels. However, with regard to extreme value analyses, small differences in the input water levels can produce large discrepancies in return water level estimates, particularly at large return periods. All water level observations are prone to natural as well as anthropogenic influences that cannot be captured by a numerical model. This causes a bias between the distributions derived from observed and modelled water levels. Thus the bias is attributed to input deficiencies e.g. resolution or scaling effects. The wind as an example has a temporal resolution of 3 hours and a spatial resolution of 2°; for simulating storm surges, this might be too coarse in order to capture all local meteorological effects.

For this reason, a bias correction method for the modelled water levels was developed; this correction almost completely eliminates the bias. The correction is performed in three steps. Firstly, tidal high water levels of observed and modelled water levels are computed and sorted in ascending order (non-parametric, empirical distribution). Secondly, the differences between the distributions of observed and modelled tidal high waters at a certain location and a specific year are calculated. The bias-correction is added to the distributions of the modelled tidal high waters, so that the bias is eliminated and the resulting values correspond to the observational tidal high waters. At each location, where observational data is available within the considered period, the bias can be eliminated using this procedure. However, as the model also generates water levels between the gauged sites, the bias-correction needs to be transferred to these locations. Thus, thirdly, the bias-

correction is interpolated to the locations between the tide gauge sites using the Inverse Distance Weighting interpolation (IDW) method. Detailed description and validation of the applied methodology is given in Arns *et al.* (2013b).

6 EXTREME VALUE STATISTICS

In this section we combine the results presented above, i.e. calculating return water levels based on the given recommendations using simulated data. Return periods and associated return water levels are calculated using both measured water level records and water level time series derived from the model hindcast after applying the bias-correction. In Figure 5a, return water levels for Hörnum are shown. As expected (Hörnum was considered for the correction), there are no differences in the estimates from the observed (blue line) and modelled (red line) water levels. Figure 5b shows the results for Pellworm Harbour (not considered for the correction but used as control station instead). In this case we find slight differences in the return water level estimates. However, up to return periods of approx. T=400 yrs, the differences are below $\Delta h \leq 2$ cm reaching $\Delta h \leq 5$ cm for a return period of T=1.000 yrs. The maximum of $\Delta h = 11$ cm occurs in T=10.000 yrs. The input period only covers 40 yrs, so that extrapolation to 10.000 yrs or even 1.000 yrs is highly debatable. The deviations referred to estimates based on observational data are therefore considered acceptable. The bias-correction is thus considered to be suitable to correct modelled water levels that are envisaged to serve as input for extreme value analyses.

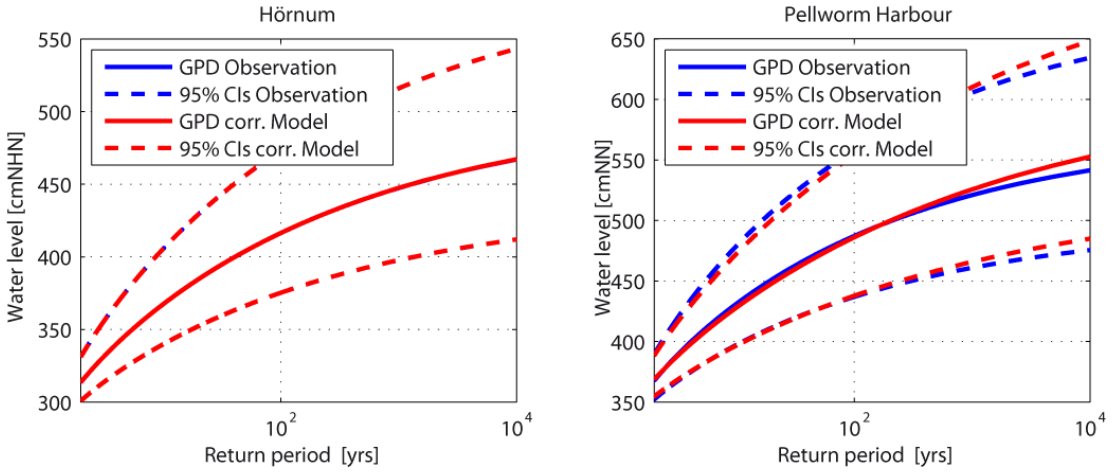


Figure 5: Return water levels for a) Hörnum and b) Pellworm Harbour

In Figure 6a, regionalized water levels with a return period of T=200 yrs are shown for the entire coastline of Schleswig-Holstein. Water levels in the southern parts of Schleswig-Holstein are higher than in the northern parts, most likely a result of shallow water effects and meteorological forcing. Figure 6b exemplarily shows regionalized return water levels for the Hallig Nordstrandischmoor, highlighting the benefit of the regionalization approach proposed here. There are no tide gauge measurements available in this area that could be used to calculate return water levels. Using the regionalization enables return water levels to be derived reliably for this un-gauged region. The information obtained here is used as a basis for the design of protection measures and are also useful risk analyses in un-gauged regions like the Halligen.

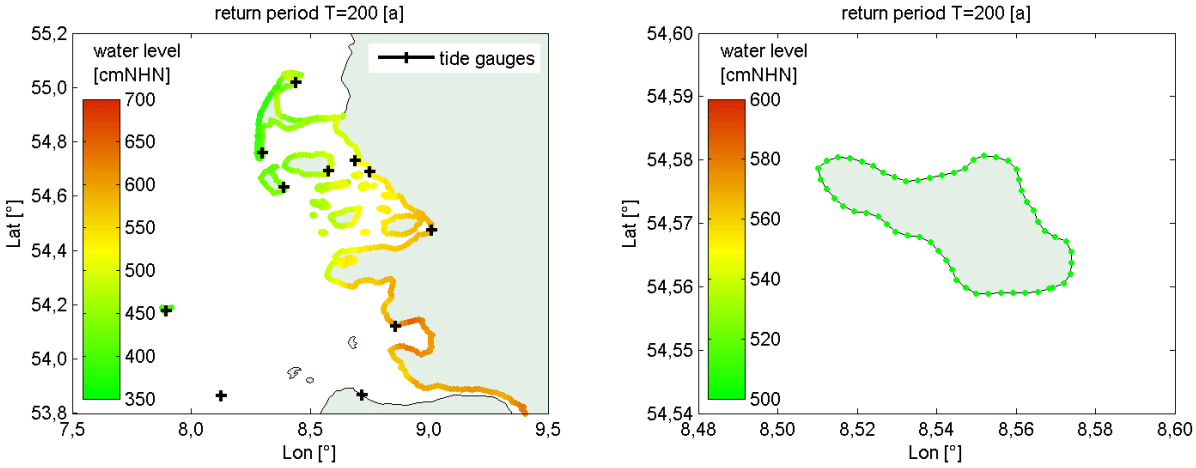


Figure 6: a) Regionalized return water levels along the coastline of Schleswig-Holstein; b) regionalized return water levels at the Hallig Hooge

7 SUMMARY

In this paper we described a method to derive information about the likelihood of extreme water levels at sites with no or little water level data, and applied this along the entire coastline of Schleswig-Holstein in northern Germany. We showed that water levels derived from a hydrodynamic model can be used to calculate reliable extreme water return levels. Especially regions with no or only few tide gauge stations can benefit from this methodology. However, a precondition is to adequately correct the bias that is generated with the numerical simulations. The bias-correction is performed first at each individual station where water level observations exist. Then the correction is transferred to the neighbouring grid points using the IDW method. As a result, regionalized extreme water return levels at un-gauged sites are obtained, that account for locally confined coastal attributes. This information can be used for planning purposes and risk analyses.

8 ACKNOWLEDGEMENTS

All analyses presented here were performed within the “ZukunftHallig” research project, a German Coastal Engineering Research Council (KFKI) project, funded by the German Federal Ministry of Education and Research BMBF through the project management of Projektträger Jülich PTJ under the grant number 03KIS093.

9 REFERENCES

Acero, F.J., García, J.A. and Gallego, M.C. (2011): *Peaks-over-threshold study of trends in extreme rainfall over the Iberian Peninsula*. Journal of Climate 24, pp. 1089–1105.

Andersen, O. B. (1995): *Global ocean tides from ERS 1 and TOPEX/POSEIDON altimetry*, Journal of Geophysical Research, Volume 100, pp. 25249-25259, doi: 10.1029/95JC01389.

Arns, A., Wahl, T., Haigh, I.D., Jensen, J. and Pattiaratchi, C. (2013a): *Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practise*, Coastal Engineering, Volume 81, pp. 51-66, http://dx.doi.org/10.1016/j.coastaleng.2013.07.003.

Arns, A., Wahl, T., Haigh, I. and Jensen, J. (2013b): *Determining extreme water return levels at un-gauged sites: a case study of the Schleswig-Holsteins coastline and Islands in north-west Germany*. Proceedings of the Coastal Dynamics conference 2013.

Bardet, L., Duluc, C.M., Rebour, V. and L’Her, J (2011): *Regional frequency analysis of extreme storm surges along the French coast*. Nat. Hazard Eartch Syst. Sci., 11, pp. 1627-1639.

Bernardara, P., Andreewsky, M. and Benoit, M. (2011): *Application of regional frequency analysis to the estimation of extreme storm surges*. J. Geophys. Res. 116.

Bocchiola et al. (2003): *Review of recent advances in index flood estimation*. Hydrology and Earth System Sciences 7, pp. 283-296.

Castellarin, A., Burn, D.H. and Brath, A. (2008): *Homogeneity testing: How homogeneous do heterogeneous cross-correlated regions seem?* Journal of Hydrology 360, pp. 67-76.

Coles, S. (2001): *An Introduction to StatisticalModeling of Extreme Values*. Springer Verlag, London.

Compo, G.P., J.S. Whitaker, P.D. Sardeshmukh, N. Matsui, R.J. Allan, X. Yin, B.E. Gleason, R.S. Vose, G. Rutledge, P. Bessemoulin, S. Brönnimann, M. Brunet, R.I. Crouthamel, A.N. Grant, P.Y. Groisman, P.D. Jones, M. Kruk, A.C. Kruger, G.J. Marshall, M. Maugeri, H.Y. Mok, Ø. Nordli, T.F. Ross, R.M. Trigo, X.L. Wang, S.D. Woodruff and S.J. Worley (2011)*: The Twentieth Century Reanalysis Projec*t. Quarterly J. Roy. Meteorol. Soc. 137, pp. 1-28, doi: 10.1002/qj.776

Cunnane, C. (1973): *A particular comparison of annual maxima and partial duration series methods of flood frequency prediction*. Journal of Hydrology 18, pp. 257–271.

Dixon, M.J. and Tawn, J.A. (1994): *Extreme sea-levels at the UK A-class sites: site-by-site analyses*. Proudman Oceanographic Laboratory Internal Document No. 65.

Dupuis, D.J. (1998): *Exceedances over high thresholds: a guide to threshold selection*. Extremes.

Environment Agency (2011): *Coastal flood boundary conditions for UK mainland and islands.* Project: SC060064/TR2: Design sea-levels. Environment Agency of England and Wales.

Haigh, I.D., Wijeratne, E.M.S., MacPherson, L.R., Pattiaratchi, C.B. and George, S. (2013a): *Estimating present day extreme total water level exceedance probabilities around the coastline of Australia*: *tides, extra-tropical storm surges and mean sea level.* Climate Dynamics. http://dx.doi.org/10.1007/s00382-012-1652-1.

Haigh, I.D., MacPherson, L.R., Mason, M.S., Wijeratne, E.M.S., Pattiaratchi, C.B., Crompton, R.P. and George, S. (2013b): *Estimating present day extreme total water level exceedance probabilities around the coastline of Australia: tropical cyclone induced storm surges*. Climate Dynamics. http://dx.doi.org/10.1007/s00382-012-1653-0.

Hosking, J.R.M., Wallis, J.R. (1993): *Some statistics useful in regional frequency analysis*, Water resources research, vol. 29, pp. 271-281.

Hosking, J.R.M. and Wallis, J.R. (1997): *Regional frequency analysis*, Cambridge University Press, Cambridge.

Jensen, J. and Müller-Navarra, S. (2008): *Storm surges on the German Coast*. Die Küste, Heft 74.

Jensen, J. (1984): *Änderungen der mittleren Tidewasserstände an der Nordseeküste.* Mitteilungen Leichtweiß-Institut der TU Braunschweig, Heft 83.

Jensen, J., Mügge, H.E. and Schönfeld, W. (1992): *Analyse der Wasserstandsentwicklung und Tidedynamik in der Deutschen Bucht*. Die Küste 53, pp.211–275 (in German).

Lang, M., Ouarda, T.B.M.J. and Bobée, B. (1999): *Towards operational guidelines for over threshold modeling.* Journal of Hydrology 225, pp. 103–117.

Mai, C.V., Van Gelder, P. and Vrijing, H. (2006): *Statistical methods to estimate extreme quantile values of the sea data*. Proceedings of the Fifth International Sysposium on Evironmental Hydraulics (ISEH-V), December 4-7, 2007 Tempe, Arizona, USA.

Merz, R. and Blöschl, G. (2005): *Flood frequency regionalisation – spatial proximity vs. catchment attributes*. Journal of Hydrology 302 (1–4), pp 283-306.

Rao, A.R. and Hamed, K.H. (2000): *Flood frequency analysis*. CRC Press, Boca Raton.

Smith, R.L. (1986): *Extreme value theory based on the r largest annual events*. Journal of Hydrology 86, pp. 27–43.

Smith, R.L., Weissman, I. (1994): *Estimating the extremal index*. Journal of the Royal Statistical Society: Series B: Methodological 56, pp. 515–528.

Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E. (1993): *Frequency Analysis of Extreme Events*, in “Handbook of Hydrology”, ed. D.R. Maidment, McGraw-Hill, New-York, NY, pp.18.1-18.66.

Tawn, J.A. (1988): *Bivariate extreme value theory: models and estimation*. Biometrika 75, pp. 397–415.

Van Gelder, P. and Nykov, M.N. (1998): *Regional frequency analysis of extreme water levels along the Dutch coast using L-moments: A preliminary study*. Stochastic models of hydrological processes and their applications to problems of environmental preservation, pp. 14-20.

Viglione, A., Laio, F. and Claps, P. (2007): *A comparison of homogeneity tests for regional frequency analysis*. Water Resources research, 43.

Wiltshire, S.E. (1985): *Grouping basins for regional flood frequency analysis*, Hydrological Sciences Journal 30 (1), pp. 151-159.

Wiltshire, S.E. (1986): *Identification of homogeneous regions for flood frequency analysis*. J. Hydrol 84, pp. 287–302.

A method to identify and form homogeneous regions for regional frequency analysis of extreme skew storm surges

Jerome Weiss[1], Pietro Bernardara[2] and Michel Benoit[1]

[1]Saint-Venant Laboratory for Hydraulics & EDF R&D LNHE, Chatou, France, Email: jerome.weiss@edf.fr

[2]Saint-Venant Laboratory for Hydraulics & EDF Energy R&D UK Centre, London, UK

Abstract

The design of effective coastal protections and the defence of coastal areas from flooding require an accurate estimation of extreme marine events, such as extreme mean sea level due to tides and storm surges and extreme wave conditions. However, due to the possible scarcity of data, local statistical analyses (i.e. based on time series of water level or wave height recorded at a single location) usually do not lead to a precise estimation of return levels of interest. Regional frequency analysis (RFA) allows reducing these uncertainties, by assuming a similar extremal behaviour at the scale of a region.
In particular, RFA based on the index flood method assumes that observations at sites in a homogeneous region follow the same probability distribution up to a local index. An important step of RFA is thus the formation of homogeneous regions. In the framework of extreme marine events, regions are generally made through statistical considerations, putting aside the physical mechanisms causing extreme marine events. In this work, a method to form physically homogeneous regions by identifying typical storms footprints is proposed. This procedure, simple to implement, is based on both a criterion of propagation of storms and a clustering algorithm. An application of this method is done on a database of extreme skew storm surges collected at 67 sites located along the Spanish, French and UK coasts.

1 INTRODUCTION

The management of extreme marine hazards is a crucial task in coastal engineering. The design of effective coastal protections and the defence of coastal areas from flooding require an accurate estimation of extreme marine events, such as extreme mean sea level due to tides and storm surges (the latter being driven by meteorological conditions and local bathymetry).

These extreme events are traditionally estimated by a local statistical analysis, from observations collected at a given site. However, durations of observations are generally too short to precisely estimate return levels of interest. These uncertainties can be reduced with regional frequency analysis (RFA), based on the index flood method developed by (Dalrymple, 1960), which tries to exploit the similarities between sites. In particular, RFA assumes that within a homogeneous region, extreme events are drawn from a common regional distribution, up to a local index representing local specificities and frequently taken as the mean event.

The grouping of sites into homogeneous regions defines the way to exploit regional information and can have a significant impact on final results. Numerous hydrological papers tried to address the formation of homogeneous regions, for example from the study of variables physically related to the phenomenon of interest. For example, (Gabriele & Chiaravalloti, 2013) recommended the use of meteorological information to form homogeneous regions in order to perform RFA of rainfall.

However, a bibliographical review indicates first that no specific method to marine hazards was developed to delineate homogeneous regions. Second, the grouping of sites is mainly done through statistical arguments and thereby excluding physical considerations. For example, (Bernardara *et al.*, 2011) estimated extreme storm surges for 18 sites located on the French coasts. The whole area was taken as homogeneous, according to a statistical test of regional homogeneity.

(Van Gelder & Neykov, 1998) performed a RFA of extreme sea levels in an area of 13 sites along the Netherlands coasts, and showed this region was statistically heterogeneous. The authors explained it by different physical processes affecting the sites (areas protected by islands, open seas and estuarine areas). The actual link between physics and statistical homogeneity is not studied and we can reasonably suppose that statistical heterogeneity may be the consequence of a physical heterogeneity.

In order to perform RFA of extreme skew storm surges, we propose a physically-based method to identify and form homogeneous regions, related to the storms yielding extremes. The typical storm footprints are identified through a clustering algorithm derived from a criterion of storm propagation. Sites are then grouped into the different regions representing storm footprints. These regions can be considered as physically homogeneous: indeed, sites from a given region are likely to be impacted by the same storms, and any storm impacting a region is likely to remain enclosed in this region. Although this method is physically-based, it is simple to implement since as it does not involve any other information than the time of occurrence of extremes.

The detail of the methodology is described in section 2. We present in section 3 an application of this method on a database of extreme skew storm surges collected at 67 sites located along the Spanish, French and UK coasts.

2 METHODOLOGY

The objective is to form physically homogeneous regions for RFA of extreme skew storm surges. The proposed method is based on the detection of typical storms footprints.

2.1 Storm extraction

A storm is here defined as a physical event generating extreme skew storm surges in at least one site in the study area. For a given site, we characterize an observation as extreme if it exceeds q_p, the p-quantile of the initial at-site skew surge series. A site is thus impacted by a storm if q_p is exceeded. Note that, following this definition, a storm is a purely statistical object, providing information on the spatial extent of the extremes generated. It indicates for each site if an extreme is observed given that at least one site is impacted by this storm.

As storms propagate in space and time, their detection is based on a spatio-temporal declustering procedure. The principle is that extremes neighbors in space and time are supposed to be part of the same storm. More precisely, two extremes are spatio-temporal neighbors if *i*) they occurred within Δ hours and *ii*) they are among the γ-nearest neighbors of each other. This spatio-temporal neighborhood relationship needs to be carefully defined, for example according to the spatio-temporal resolution of observations, the possible missing values and the physical propagation of the considered phenomenon. Thus, three parameters are required to detect a storm: p, setting its impact on a given site, and (Δ, γ) which are related to its spatio-temporal propagation. (p, Δ, γ) should be chosen in order to guarantee a proper detection of these physical events.

We consider here the hypothesis, often accepted in the literature, that the declustering procedure is leading to a sample of independent storms. Moreover, at a given site, one or more extremes can occur during a same storm, according to its duration. When several extremes appear, only the maximum value is retained to get independent extremes at site scale.

2.2 Formation of physically homogeneous regions

We attempt in this section to partition the set of sites so that each resulting group represents a typical storm footprint. If N sites compose the study area where T storms are observed, let Z_t^i be the Bernoulli variable, where $Z_t^i = 1$ if site i is impacted by the storm $t = 1,\ldots,T$. A criterion of storm propagation $p_{i,j}$ is defined as the probability that both sites i and j are impacted by a storm given that one of them is:

$$p_{i,j} = P\left(Z_t^i = 1, Z_t^j = 1 \mid Z_t^i + Z_t^j \geq 1\right) \quad (1)$$

For each pair of sites (*i*, *j*), these probabilities are estimated from the observed storms. Then, a dissimilarity index $d_{i,j} = 1 - p_{i,j}$ is defined and computed for each pair of sites. In particular, $d_{i,j} = 0$ ($d_{i,j} = 1$) indicates that any storm impacting *i* or *j* necessarily hits (avoids) the other.

The next step is to group all sites into *R* disjoint regions, according to their similarity in terms of the criterion of storm propagation (1). By definition, the resulting partition can be considered to represent storms footprints. We use Ward's hierarchical clustering algorithm (Ward, 1963); this agglomerative hierarchical method initially assigns each site to its own region, and the closest pair of regions is merged until there is only one region. A dendogram can be used to represent the hierarchy of regions.

Thus, for fixed *R*, the study area is divided into *R* regions. Several configurations of storms footprints can be obtained when *R* varies. In order to determine the optimal value for *R*, the evolution of the standardized dendrogram heights in function of the number of clusters is examined. This is based on the fact that a break indicates that two dissimilar clusters have been merged (Martinez & Martinez, 2004). The most relevant storms footprints can be identified with this procedure. Statistically speaking, it means that a storm impacting a given region is likely to remain enclosed in this region, and sites in this region are likely to be impacted by the same storms.

2.3 Regional statistical homogeneity

RFA of extreme storm surges requires statistically homogeneous regions. The physically homogeneous regions obtained from section 2.2 should then be checked to be also statistically homogeneous.

The storms previously extracted in section 2.1 represent physical events generating extremes and are used to form physically homogeneous regions. However, for statistical aspects, a sub-selection of these storms is proceeded in order to focus on the most intense events. In particular, we redefine storms in such a way that there is $\lambda = 1$ storm a year on average at each site. So, this procedure leads to the observation of n_i extreme skew storm surges at site *i*, sampled from the random variable $X^i X^i \sim \mathrm{GEV}(\xi_i, \alpha_i, k_i)$.

Discordant sites can be identified through the discordancy criterion *D* of (Hosking & Wallis, 1997). It measures if a given site is significantly different from the other sites in the region, in terms of L-moments. A site can be declared discordant if $D > 3$. Besides, the statistical homogeneity of a region can be assessed with the Hosking and Wallis test (Hosking & Wallis, 1997). Their heterogeneity measure *H* informs if the observed dispersion between sites is comparable to what would be expected in a statistically homogeneous region. In particular, the region can be considered as statistically homogeneous if $H < 2$, and heterogeneous otherwise.

For each storm footprint, the following procedure is applied:

- *i)* Computation of the heterogeneity measure *H*. If $H < 2$ then go to *iv*), else go to *ii*).
- *ii)* Computation of the discordancy measures D. If no site is discordant then go to *iii*), else remove the sites with $D > 3$ and compute a new heterogeneity measure *H'*. If $H' < 2$ then go to *iv*), else go to *iii*).
- *iii)* Subdivision of the region into $R = 2$ new storms footprints (section 2.2). For each sub-region, go to *i*).
- *iv)* The region, both physically and statistically homogeneous, is used for RFA.

RFA can be then performed with these regions, both physically and statistically homogeneous.

2.4 Estimation of the regional distribution (method of Hosking & Wallis)

For a homogenous region of *N* sites, let $\hat{\mu}_i = \overline{X}^i$ be the mean value of the site $i = 1,\ldots,N$. By regional homogeneity, the normalized variable $X^R = X^i / \hat{\mu}_i$ is supposed to be independent of *i*, with cdf F_R.

x_F^iAs the observations sampled from X^i are exceedances over a high threshold, the regional distribution F_R is taken as the Generalized Pareto Distribution (GPD) (Pickands, 1975). Let ξ_R, α_R and k_R be the location, the scale and the shape parameter of the GPD, respectively. The p-quantile of F_R is:

$$x_p^R = \begin{cases} \xi_R + \frac{\alpha_R}{k_R}\left(1-(1-p)^{k_R}\right), & k_R \neq 0 \\ \xi_R - \alpha_R \log(1-p), & k_R = 0 \end{cases} \quad (2)$$

The parameters of F_R can be estimated with the L-moments method (Hosking & Wallis, 1997). Let $\hat{\lambda}_r^i \hat{\lambda}_r^i$ be the estimator of the r-order L-moment for the site i and $\hat{\tau}_3^i = \hat{\lambda}_3^i / \hat{\lambda}_2^i$ $\hat{\tau}_3^i = \hat{\lambda}_3^i / \hat{\lambda}_2^i$its estimation of the L-skewness. The equivalent regional quantities are $\hat{\lambda}_r^R = \sum_{i=1}^N n_i \left(\hat{\lambda}_r^i / \hat{\mu}_i\right) / \sum_{i=1}^N n_i$ and $\hat{\tau}_3^R = \sum_{i=1}^N n_i \hat{\tau}_3^i / \sum_{i=1}^N n_i$. The p-quantile of the site i is then estimated by $\hat{x}_p^i = \hat{\mu}_i \hat{x}_p^R$, with:

$$\hat{k}_R = \left(1-3\hat{\tau}_3^R\right)/\left(1+\hat{\tau}_3^R\right), \ \hat{\alpha}_R = \left(1+\hat{k}_R\right)\left(2+\hat{k}_R\right)\hat{\lambda}_2^R, \ \hat{\xi}_R = \hat{\lambda}_1^R - \hat{\alpha}_R/\left(1+\hat{k}_R\right) \quad (3)$$

In particular, the regional T-year return level is estimated by $\hat{x}_{1-\frac{1}{\lambda T}}^R$; for site i, the local T-year return level is estimated by $\hat{x}_{1-\frac{1}{\lambda T}}^i$.

3 APPLICATION

3.1 Skew storm surge data

The raw data used in this study are temporal series of hourly sea level observations collected at 67 ports along the Spanish, French and U.K. coasts (Figure 1a). French data are supplied by SHOM (Service Hydrographique et Océanographique de la Marine, France) and available on the REFMAR (Réseaux de référence des observations marégraphiques) website, while Spanish and UK data are respectively supplied by IEO (Instituto Español de Oceanografía, Spain) and BODC (British Oceanographic Data Centre, UK). The series range from 1846 (Brest, France) to 2011, show a mean effective duration of 31 years and display missing values.

Local mean sea levels may be affected by eustatism (i.e. the long-term alteration of mean sea levels caused by either geological or climate changes), while tidal predictions are given for the present time. In order to calculate the actual surges of past periods, the sea level must be corrected from a possible eustatism. If annual sea levels, calculated following the PSMSL (Permanent Service for Mean Sea Level) recommendations, show significant linear trends, then raw sea level data are corrected to ensure the stationarity of annual sea levels.

In regions with strong tidal influence, coastal flooding hazard is more marked around the times of high tide. Therefore, we restricted our attention to skew surge series, in order to describe the surge contribution at the maximum tidal level. The skew surge is defined as the (algebraic) difference between the maximum observed sea level around the time of theoretical (predicted) high tide and the predicted high tide level. Thus, the resulting skew surge series have a temporal resolution of approximately 12.4 hours. For a more detailed introduction on skew surges, see (Bernardara *et al.*, 2011; Simon, 2007; Weiss *et al.*, 2012).

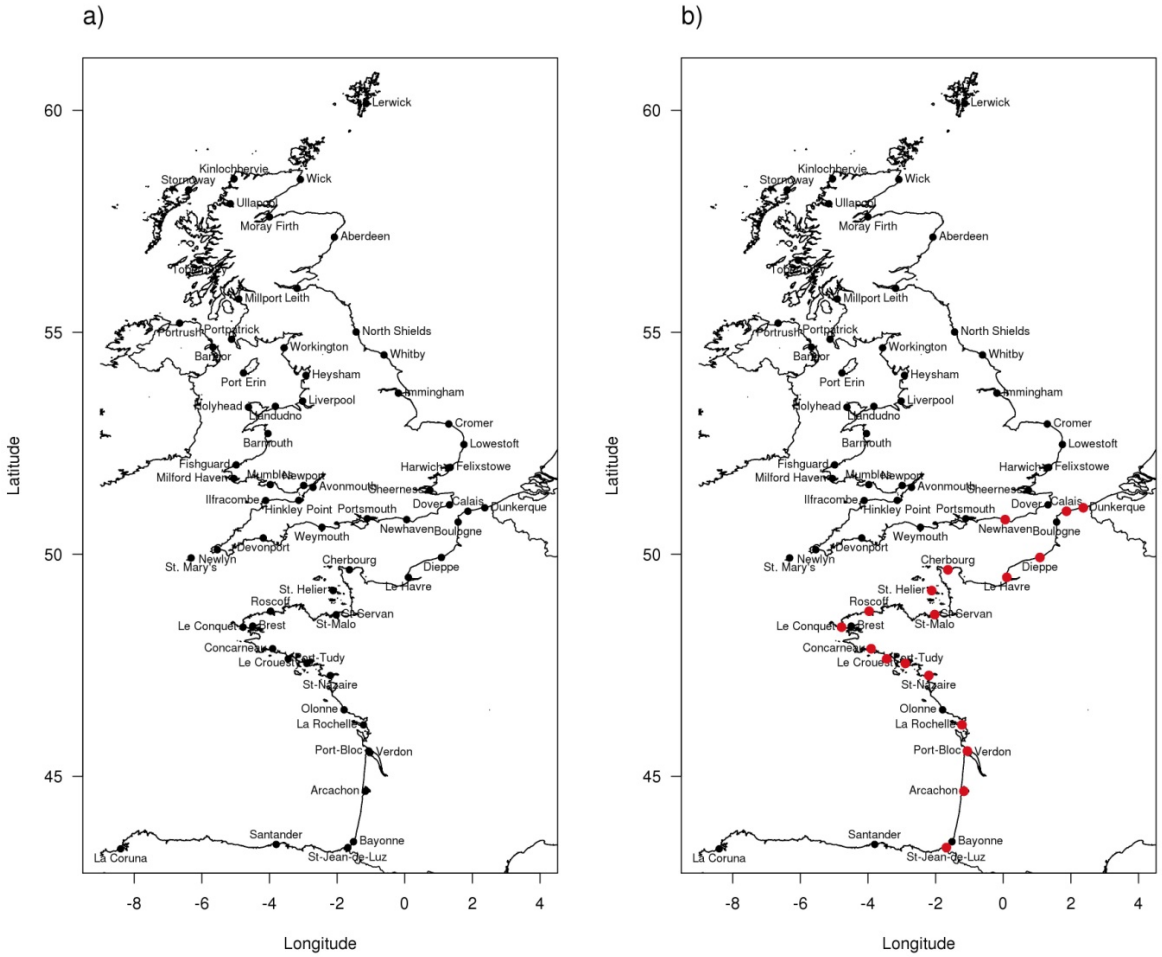


Figure 1: a) Location of the 67 sites. b) Xynthia storm of February 2010 (red dots indicate impacted sites)

3.2 Formation of homogeneous regions

The parameters (p, Δ, γ) required to detect a storm (section 2.1) are set in order to faithfully reproduce the dynamics of the most intense storms present in the database, while taking into account the spatio-temporal resolution of observations and the possible missing values. A sensitivity analysis was performed by considering especially the North Sea Flood of 1953, the Great Storm of 1987, Martin (1999) and Xynthia (2010). The configuration (p = 0.995, Δ = 24 hours, γ = 14) was retained, leading to the extraction of 1706 storms. The Xynthia storm of February 2010 is shown in Figure 1b.

The criterion of storm propagation (1) is then estimated for each pair of sites, from the 1706 storms. The Ward's hierarchical classification, applied on these quantities, is represented by the dendrogram in Figure 2. The evolution of its standardized heights (not shown) reveals that a partition into 4 clusters could correspond to the most typical storm footprints. These 4 physically homogeneous regions are shown in Figure 3a: the Bay of Biscay and the English Channel (region 1), the North Sea (region 2), the North UK coasts (region 3), the West UK coasts (region 4).

The procedure of section 2.3 is applied to get both physically and statistically homogeneous regions. Note that the extraction of local samples of extremes such that there is, on average, λ = 1 storm a year at each site lead to retain 800 storms among the 1706. All but region 4 are statistically homogeneous. Region 4 (H = 4.10) is subdivided into two inner storm footprints (the Irish Sea and the Bristol Channel), which are statistically homogeneous. The 5 resulting homogeneous regions are shown in Figure 3b. It should be noted that the whole area is not statistically homogeneous (H = 3.22), underlining the interest of a subdivision into smaller regions. Finally, the proposed method allows the increase of the overall statistical homogeneity.

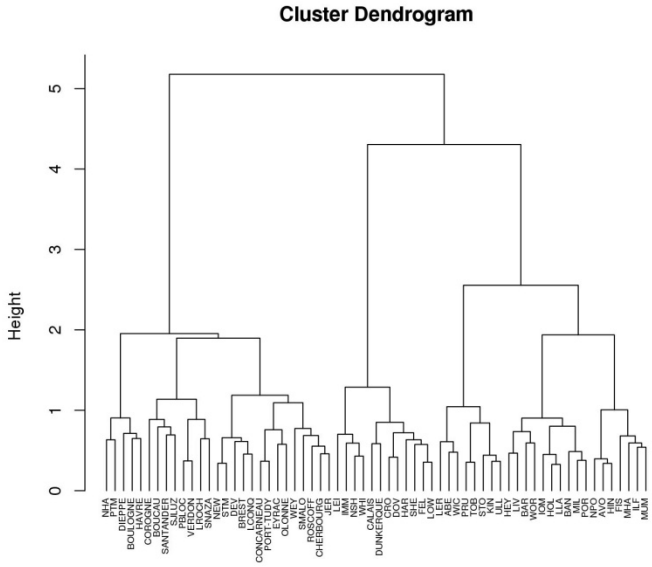


Figure 2: Hierarchical clustering from the criterion of storm propagation (dendrogram)

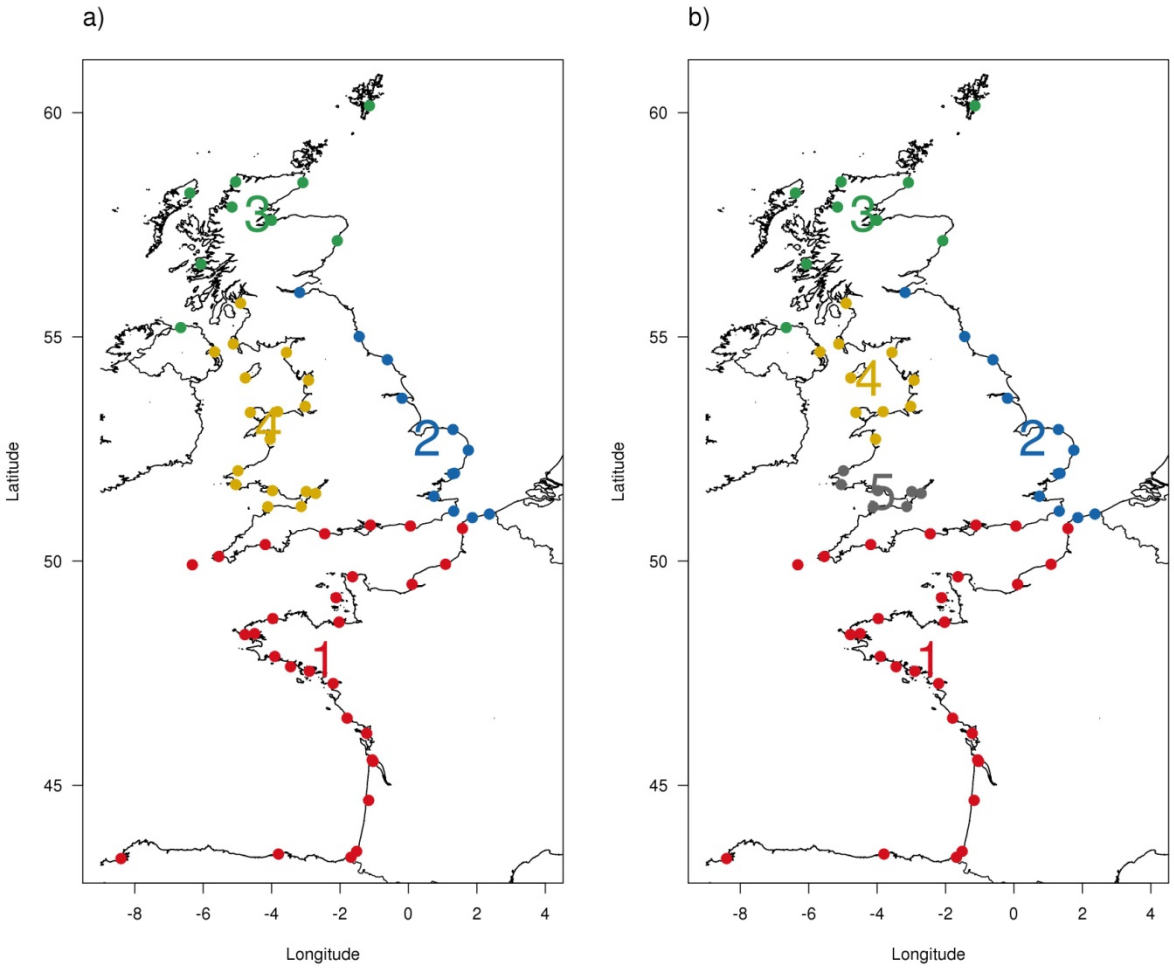


Figure 3: a) Physically homogeneous regions. b) Physically and statistically homogeneous regions

3.3 Estimation of extreme skew storm surges

The method of Hosking and Wallis (section 2.4) is used to estimate the parameters of the regional GPD. From Table 1, negative values of the shape parameter are obtained for regions 1 and 5, suggesting a higher intensity of extremes (unbounded GPD). The corresponding regional return level plots are represented in Figure 4, where the 95 % confidence intervals are obtained by bootstrapping the observed storms.

Table 1: Parameters of the regional GPD

Region	ξ_R	α_R	k_R
1	0.845	0.153	-0.014
2	0.835	0.176	0.073
3	0.877	0.123	0.002
4	0.834	0.201	0.214
5	0.808	0.182	-0.048

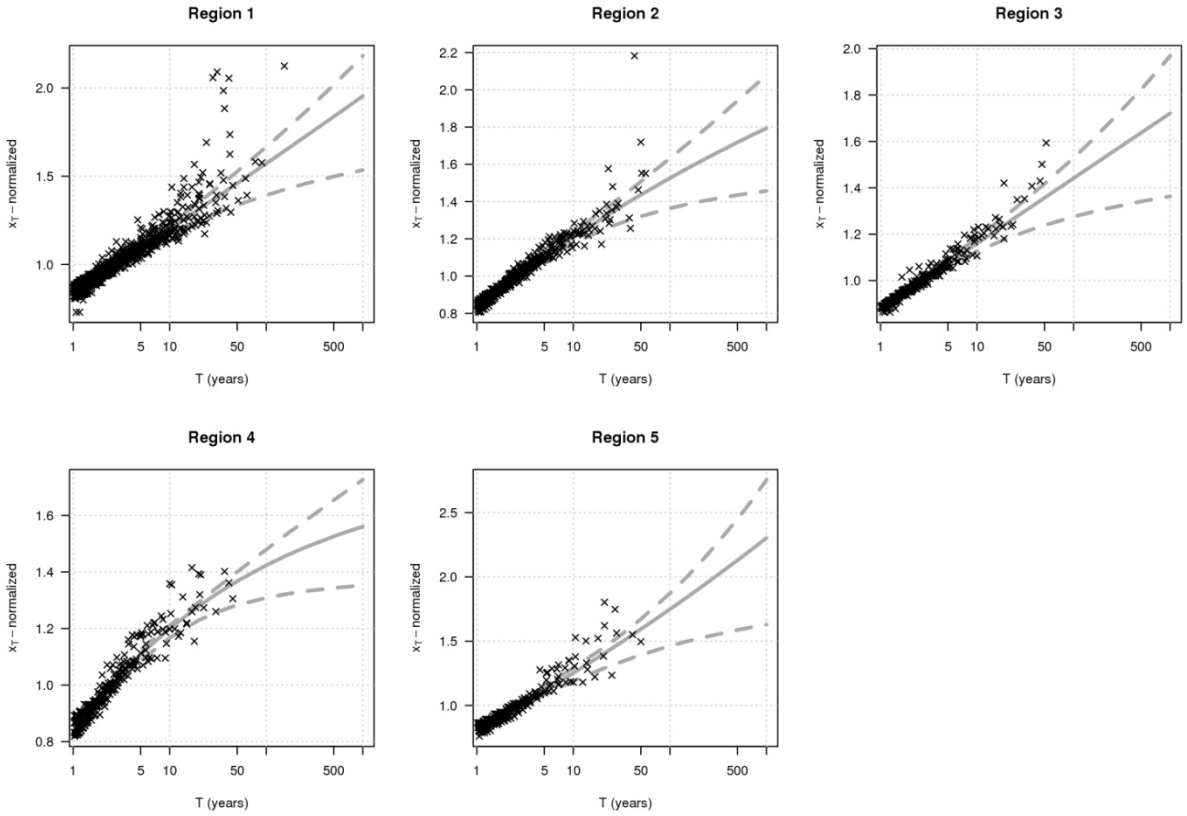


Figure 4: Regional return level plots with 95% bootstrap confidence intervals (crosses represent normalized observations from each site in the region)

4 CONCLUSION

With the aim to estimate extreme skew storm surges, a possible solution to reduce uncertainties inherent to traditional local statistical analyses is RFA, which uses the common regional information shared by similar sites.

RFA requires to group sites into homogeneous regions. However, in the context of extreme marine events, this crucial step still remains an open question. We here propose a method to identify and form physically homogeneous regions by finding typical storms footprints. The procedure, simple to

implement, relies on a clustering algorithm based on a criterion of propagation of storms. Specifically, extreme skew storm surges from a given region are likely to be generated by the same storms, and any storm impacting a region is likely to remain enclosed in this region.

The proposed method is used to form homogeneous regions from a database of skew storm surges collected at 67 sites located along the Spanish, French and UK coasts. The whole area, statistically heterogeneous, needs to be refined. 5 regions, both physically and statistically homogeneous, are identified as typical storms footprints: the Bay of Biscay and the English Channel, the North Sea, the North UK coasts, the West UK coasts. It appears that this physically-based delineation of homogeneous regions leads towards an increase of the overall statistical homogeneity. A RFA is then performed on these regions to estimate extreme skew storm surges in the study area, and, at the same time, to highlight regional differences in their probabilistic behavior.

5 ACKNOWLEDGMENTS

The permission to publish the results of this ongoing research study was granted by the EDF (Electricité De France). The results in this paper should, of course, be considered as R&D exercises without any significance or embedded commitments upon the real behavior of the EDF power facilities or its regulatory control and licensing. The authors would like to acknowledge the SHOM (Service Hydrographique et Océanographique de la Marine, France), the REFMAR (Réseaux de référence des observations marégraphiques, France), the BODC (British Oceanographic Data Centre, UK) and the IEO (Instituto Español de Oceanografía, Spain) for providing the data used in this study.

6 REFERENCES

Bernardara P., Andreewsky M. and Benoit M. (2011): *Application of the Regional Frequency Analysis to the estimation of extreme storm surges*, J. Geophys. Res.116, C02008, pp. 1–11.

Darlymple, T. (1960): *Flood Frequency Analysis*, 1543–A, US Geological Survey, Water Supply Paper.

Gabriele, S. and Chiaravalloti, F. (2013): *Using the Meteorological Information for the Regional Rainfall Frequency Analysis: An Application to Sicily*, Water resources management 27 (6), pp.1721–1735.

Hosking J. R. M. and Wallis J. R. (1997): *Regional Frequency Analysis. An approach based on L-moments*, Cambridge, Cambridge University Press.

Martinez W. L. and Martinez A. R. (2004): *Exploratory Data Analysis with MATLAB*, Chapman and Hall/CRC.

Pickands, J. (1975): *Statistical Inference Using Extreme Order Statistics*, The Annals of Statistics 3 (1), pp.119–131.

Simon, B. (2007): *La marée océanique côtière*, Editions de l'Institut Océanographique.

Van Gelder P. H. A. J. M. and Neykov N. M. (1998): *Regional Frequency Analysis of Extreme Water Level along the Dutch Coast using L-moments: a preliminary study. Preprints of the International Scientific Conference on Stochastic models of hydrological processes and their application to problems of environmental preservation*, Moscow, Russia.

Ward, J. (1963): *Hierarchical grouping to optimize an objective function*. J. Amer. Stat. Assoc. 58, pp. 236–244.

Weiss J., Bernardara P., Andreewsky M. and Benoit M. (2012): *Seasonal autoregressive modeling of a skew storm surge series*, Ocean Modelling 47, pp.41–54.

Simulated future tides and sea state in the Elbe estuary

Hartmut Hein[1], Stephan Mai[1] and Ulrich Barjenbruch[1]

[1]German Federal Institute of Hydrology, Koblenz, Germany, Email: hein@bafg.de

Abstract

Long-term simulations of the tides and the sea state in the Elbe estuary are discussed for expected future conditions in relation to present-day conditions and multi-scale variability. Simulations of the tides (1950 – 2100) are conducted with a three-dimensional limited area model of the Elbe estuary, which is offline nested into a baroclinic circulation model of the North Sea. For determining the sea state a statistical approach links the statistics of the water level, the wind speed and direction to the statistics of wave parameters, like significant wave height, wave period and wave direction. The statistical transfer uses results of numerical simulations of wave propagation.

For one typical point in the river mouth monthly maxima, yearly maxima, 19 yearly maxima and the maxima of 100 years are calculated for tidal water levels as well as for significant wave heights. They are statistically analysed afterwards. Two modern methods [detrended fluctuation analysis (DFA) and wavelet analysis (WA) are applied to estimate the temporal correlations of the numerical long term simulations, for both monthly means and maxima.

Process based downscaling of a global climate model into an estuary is a quite well functional method to estimate future changes of mean conditions and maxima – if an uncertainty analysis of the results is done. Non-stationary and multi-decadal hydrodynamic responses of estuaries to climate change can be estimated. A critical challenge in supporting adaptation is the linkage between vulnerability research and coastal management decisions with respect to multi-scale variability.

1 INTRODUCTION

In coastal regions and estuaries physical processes influence many economic, ecologic processes and also security issues. Global climate change has a high potential to influence both the persistence and the transport pathways of water masses and its constituents in tidal waters and estuaries (Dietrich *et al.*, 2013).

Sensitivity studies (e.g. Mai *et al.*, 2004) show the variation of tidal water levels, of significant wave heights and the morphology at the North Sea coast as the result of climate change. In the long term context of climate change, these physical processes are subject to changes, too (Hein et al., 2011b). In order to get the impression of the future changes and the probability of their occurrence, physically consistent long term simulations are needed to describe how wind waves and currents interact and control processes like erosion, sedimentation and biological production.

It is widely accepted, that the climate-related sea level rise (SLR) influences the long-term coastal processes. In this context the term “climate mean” is defined as the characteristic frequency distribution of local conditions and processes for a sufficient period of time. This period reflects the probability density of states and processes of the typical conditions in the region, which is in this study the mouth of the river Elbe. An almost linear secular rise of about 1-2 mm per year (e.g. Wahl *et al.*, 2010; Hein *et al.,* 2011c) has already been observed in the southern German Bight. The future acceleration of global SLR is expected (IPCC, 2007), historic acceleration for the German Bight is not to be found significant (Hein *et al.*, 2010).

Variability of the SLR is acts on almost all scales. Several of the sales are representative for typical atmospheric timescales of months to several years (Dangensdorf, 2013). The SLR is positively correlated with the changes of the NAO on time-scales of 4 to 7 years (Hein, 2011b; Dangendorf,

2012). Variability on scales in the range of the Nodaltide is indicated since 1930 and before 1900; periods in order of 30 to 40 years are important, their amplitudes increase with time. Additionally, periods of approximately 60 to 80 years are present in the sea level of the German Bight (Hein *et al.*, 2011b).

The expected future changes of the global sea level in the 21st century are mainly determined by the steric expansion of the ocean due to global warming. Additionally increasing fresh water supply from melting of the two ice sheets over Greenland and the Antarctic and from inland deglaciation accelerates the SLR in the 21st century. However, the regional sea level rise must be determined by the regional distribution of globally added melt water masses due to gravitational effects and also by barotropic and baroclinic ocean dynamics due to changing density distributions (Mathis, 2013). In the Elbe estuary the glacial isostatic adjustment causes land subsidence in order of 5 cm to 10 cm (Hein *et al.*, 2011c). For this model study, we use the approach from Mathis (2013), who implements the sea level rise at the out boundary of the North Sea model in form of a scenario. Based on this scenario this study estimates monthly maxima, yearly maxima and the maxima of 19 years. Both, tidal water levels and significant wave heights are calculated and statistically analysed.

Tide and wave climate forecasting is one major issue for coastal management. For sedimentation processes often not the mean states the important ones, but the maxima are in the focus of research - simply related to the quadratic law in the calculation of shear stresses. From former studies (Mudersbach *et al.*, 2013; Weisse *et al.*, 2011) it is known that the storm surges and also wave heights in the German Bight vary on long time scales but they show no significant trend. It is expected that extreme sea levels increase primarily as a result of mean sea level changes or as a result of increasing of the tidal amplitudes.

2 APPROACH

2.1 Simulations

The changes in the statistical maxima of water levels and wave parameter, like monthly maximum high water and monthly maximum significant wave height (Hs) are derived by the use of long-term regionalized coupled numerical modeling of atmosphere and ocean (Hein *et al.*, 2013); the so called model chain (MC) is used. The MC implemented in the research program KLIWAS of the German Federal Ministry of Transport, Building and Urban Development, by the Federal Institute of Hydrology together with several partners downscales *one* climate scenario towards long-term simulations of the German North Sea estuaries.

The overall MC starts with emission scenarios. These are used to run various global climate models, to derive atmospheric and oceanographic parameters on the global scale. It is necessary to transform the results of the global climate models with regional downscaling into results for the specific region. This is usually done with the uncoupled models of ocean and atmosphere. The last step is to scale the regional climate models down towards the certain stretch of the coastline. The lack of tidal information, in most of the global climate models is one challenge for simulations of coastal processes. The regional topography of the simulations are shown in Figure 1.

The second missing parameter in global climate models is the SLR. This stands in contrast to the importance of the SLR for the future change of the regional tidal system. Since often climate models are volume-conserving, they cannot account for SLR due to thermal expansion. Sea level changes due to increasing fresh water supply from melting off the ice are neglected, too. For the HAMSOM simulations Mathis (2013) induce estimations of the different components of global sea level rise and continually add them onto the sea surface elevation at the open boundaries. Corresponding to the upper limit of the IPCC bandwidth, in this study SLR of about 50 cm from the period 1990-1999 to 2090-2099 is added at the North Sea boundary.

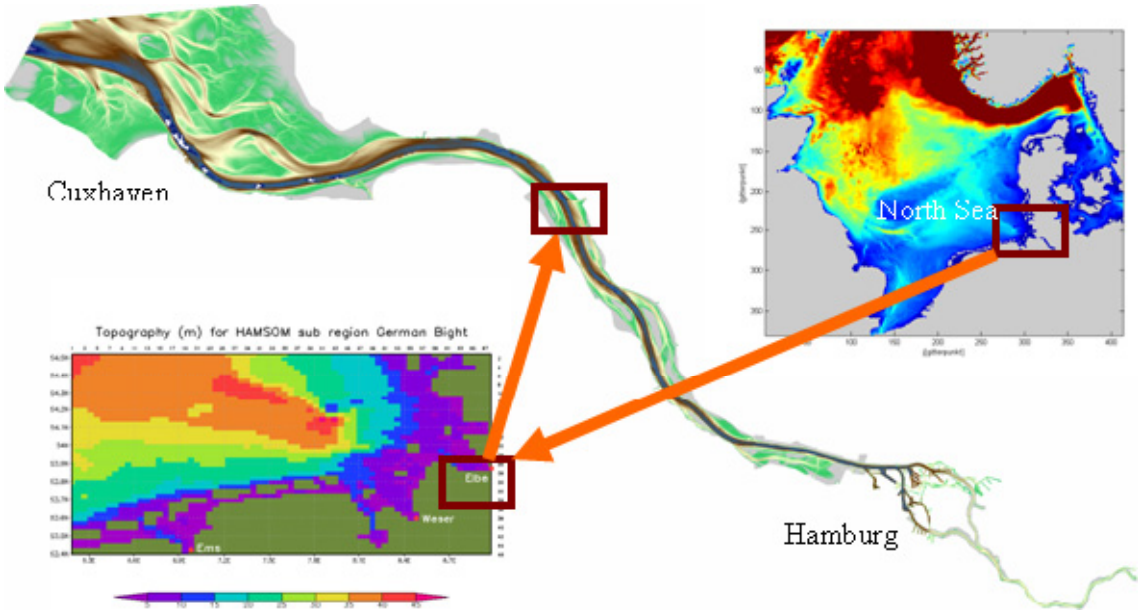


Figure 1: Regional topography used by the simulations

Time-series of water level and wind are derived from the global climate run A1B MPI-OM, which is regionalised to the North-Sea with the offline coupled models HAMSOM/Remo (Pohlmann, 2006) with the use of an additional forcing from a global tide model. To simulate the circulation and the sea-level, the hydro-numerical model HAMburg Shelf Ocean Model (HAMSOM) is used. HAMSOM was first set up in the mid-eighties by Backhaus (Backhaus, 1985). In general, it is a three-dimensional, prognostic-baroclinic, frontal- and eddy-resolving model with a free surface. The numerical scheme of HAMSOM is defined in z-coordinates on an Arakawa C-grid. The governing equations for shallow water combined with the hydrostatic assumptions are implemented. The basic equations can be found in Pohlmann (1996).

The simulation of the estuarine circulation yield several numeric requirements to the model (Hein *et al.*, 2007). Therefore, high-order formulations are used for the momentum equation and the transport equation. The importance of diffusion processes on (de-) stratification in estuaries is considered by sub-grid stochastic simulations: The vertical turbulent viscosity is calculated by a Kochergin-Pohlmann-Scheme (Pohlmann, 1996). The horizontal sub-grid processes are estimated by a Smagorinsky-Scheme (Hein, 2008).

The applicability of the regional circulation model was shown in several studies (Hein *et al.*, 2011a; Hein *et al.*, 2012; Hein *et al.*, 2013). It turns out, that the local model, despite the low resoltution transports the tidal wave to the port of Hamburg in an adequate manner. However, numerical models may be useful tools to get insight in the coastal processes of the system being modelled, but poor input data leads to uncertain model results (Spek, 2013). Especially by the use of a climate MC additional stochastic analysis should be used.

To estimate the significant wave hight (Hs) the results of the long term hydrodynamic models are combined with short-term numerical modelling of waves. A flow chart of the calculation scheme for Hs (Mai *et al.*, 2008) is given in Figure 2. For the short-term calculations of wave parameters as a function of water level, wind speed and direction the numerical model SWAN (Booij *et al.*, 1999) is used. The calculations were carried out on a curvilinear computational grid of the topography of the year 2006 with a resolution of approx. 20 m along the river and approx. 2 m across the river (Berkhahn & Mai, 2004).

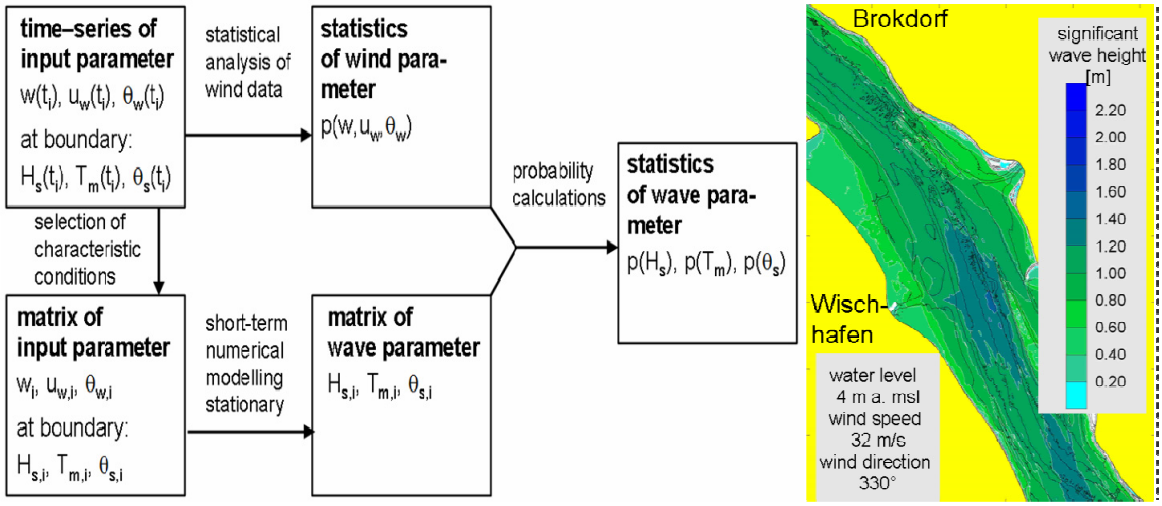


Figure 2: Calculation scheme combining time-series of water level and wind with wave parameters

The wave field is calculated for 840 combinations of different boundary conditions, i.e. water levels and winds. The set of wave simulations is used to derive transfer functions from water level and wind to wave parameters (Figure 3). Between the discrete values the spline interpolation is used. On the right hand side of Figure 3 a typical wave field in the study area can be seen.

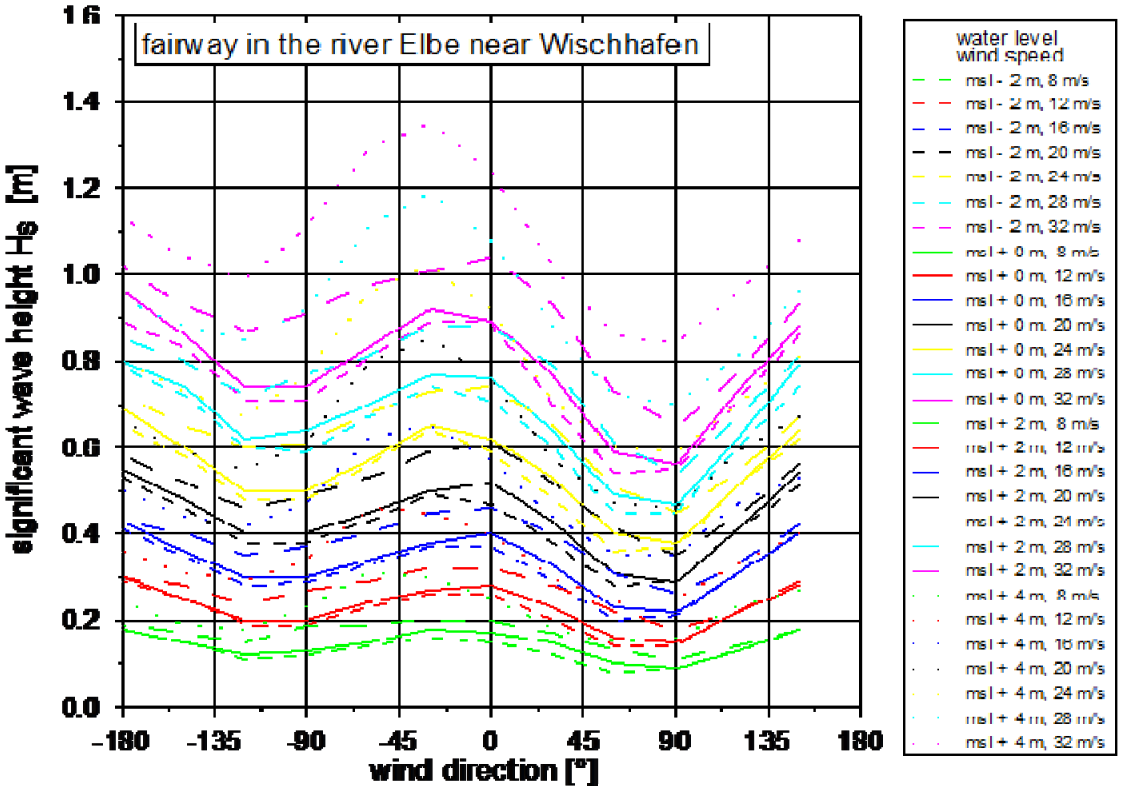


Figure 3: Transfer functions for wave height

2.2 Analysis tools

Hurst (1951) analysis of hydrological data indicates that the fluctuations in hydrology are self-similar over a wide range of time scales, with no single characteristic scale. This should be the same with mean *coastal* hydrological parameters like sea level and sea state. As one conclusion of the Hurst's work can be drawn, that it is general difficult to distinguish trends from long-term correlations. Stationary long-term correlated time series feature persistent behaviour, which may cause the

detection of erroneous trends. In the last years, several methods such as detrended fluctuation analysis (DFA, Peng *et al.*, 1994) and wavelet analysis (WA, Torrence & Compo, 1998) have been developed. They are able to determine long-term correlations in the presence of trends.

3 RESULTS AND DISCUSSION

3.1 Monthly, yearly and 19 yearly and 100 yearly Maxima

Figure 4a shows the monthly, yearly, 19 yearly and 100 yearly maxima of the long term simulations in relation to the monthly, yearly and 19 yearly mean sea level. In the mean values the added sea level rise can be seen. The rate of the rise of the maxima differs not significantly from that of the means. The variability of the maximum values is statistically reasonable higher than that of the mean values. Long-term statistical maxima (19 year maxima and 100 year maxima) of the sea level are triggered by single (storm) events. Therefore, no future changes can be predicted by the means of one member of a MC. This means that it can be expected that the SLR published by the IPCC related to the A2B Scenario cannot significantly influence statistical maxima. The relation of the SLR to the powerfulness of one single storm event is low. However, the monthly and yearly maxima underlie a slightly higher mean rise than the mean sea level. This result is important information regarding sedimentation processes or ecologic proposes because these short time maxima may be relevant for these processes.

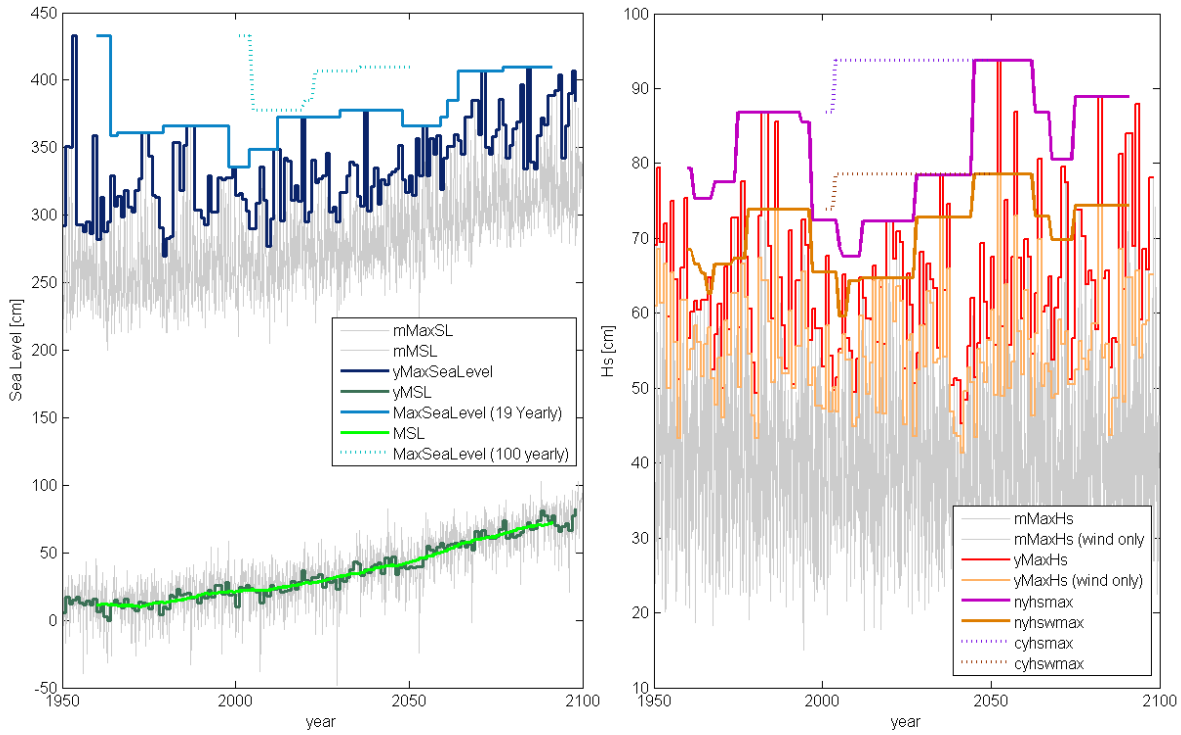


Figure 4: Monthly, yearly, 19 yearly and 100 yearly maxima and mean of sea level (left) and significant wave heights (right)

Figure 4b shows the monthly, yearly and 19 yearly maxima of the long term simulations of the significant wave height (Hs). Two simulations are compared, one including sea level data, the second with a constant sea level with time. If the sea level is included the maxima are slightly higher for all time scales. A clear trend is not to be seen in the monthly mean, nor in the yearly and nor in the 100 yearly maxima. If the difference between the two simulations is taken into account, it seems that as a result of the SLR the 19 yearly maximum increases with time. However, the clear stochastically statement suffers, due to the use of one member of a MC.

3.2 Persistence of Monthly sea levels

By use of the DFA we calculated the Hurst coefficient with the value 1.05 for the monthly mean sea level and 0.88 for the monthly maximum sea level. However, monotonic trends tend to result in an over-estimation of the Hurst exponent and uncorrelated data superimposed on a long-term trend will exhibit autocorrelation (Bhatthacharya *et al.*, 1983). Our time series form the simulations underlie the before given trend in the mean sea level. Therefore we calculated the Hurst coefficient again, based on detrended time series. We calculated the Hurst coefficient for the detrended time series with the value 0.85 for the monthly mean sea level and 0.75 for the monthly maximum sea level.

In detail figure 5 demonstrates that there is a relatively higher intra yearly Hurst coefficient, which indicates the seasonal cycle. The coefficient changes more to white noise in the short term variability of 1 to 4 years. On longer scales the persistence turning more to pink noise which can typically be estimated on such time scales for atmospheric-oceanographic components (Fernández *et al.*, 2003). The results of the detrending of the time series are mostly visible on the longer scales.

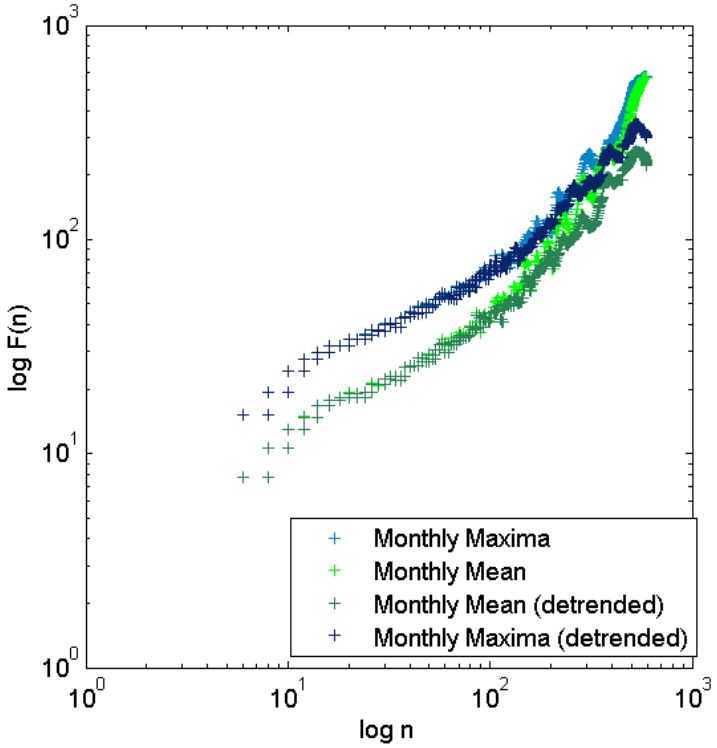


Figure 5: DFA from the monthly mean and maximum sea level

The calculated mean values for the Hurst coefficient are typical for hydrological regimes, which variability is suggested to be somewhere between white noise and pink (1/f) noise. For example Sakalauskienė (2003) calculated for the Nemunas river (Lithuania) the value of 0.67 for the Hurst coefficient. It might be surprising that also the monthly maxima underlie long term persistence. However, Lye & Lin (1994) also estimated the persistence in the yearly maxima of several Canadian rivers.
Our literature research indicates that in coastal sciences the calculation of persistence is not common. However, our result indicates that coastal time-series of sea level are dominate by long term multi-scale variability which complicates the trend estimation and hinders us estimating the acceleration of the SLR.

3.3 Persistence of Monthly significant wave heights

By use of the DFA we calculated the Hurst coefficient with the value 0.4 for the monthly mean Hs and 0.47 for the monthly maximum Hs if considering the sea level variations. For the simulated Hs forced by wind only the DFA results in a Hurst coefficient of 0.4 for the monthly mean Hs and 0.46

for the monthly maximum Hs. Hence the persistence remains the same, which demonstrate that the sea level does not influence the persistence of Hs on a significant level. The results indicate that the significant wave heights are dominated by white noise processes.

In detail, it can be see in Figure 6 that only on intra-annual scales, increased Hurst coefficients are to be found, which indicate the seasonal cycle. Persistence on higher time scales could not be seen. However, our model chain is forced by a global climate model. Bakker & van den Hurk (2012) analyzed the sea level pressure in the North Atlantic region. They estimate a Hurst coefficient in the Range of 0.58 to 0.74 from observations and almost white noise from climate models. So it may be concluded that there should also persistence in the significant wave heights if the estimations base on observations.

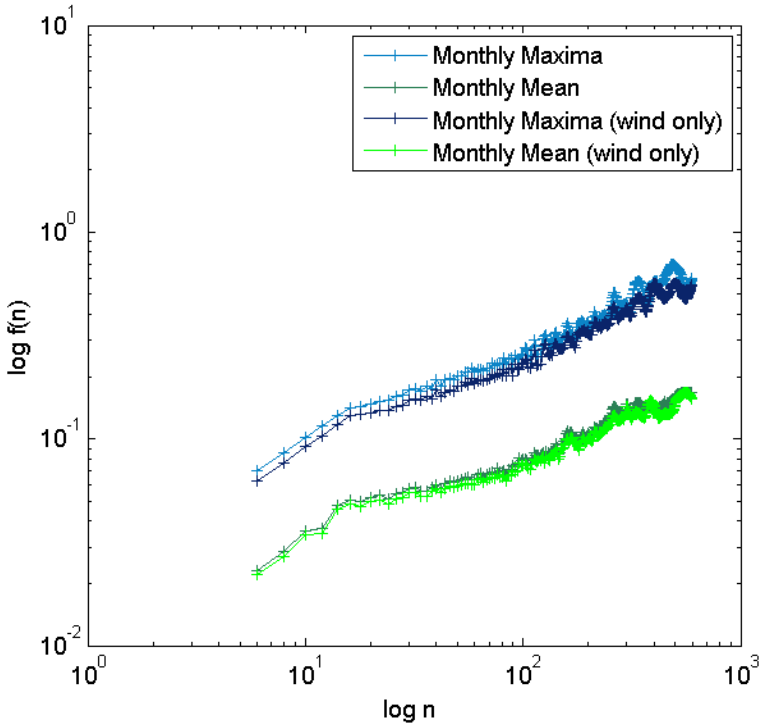


Figure 6: DFA from the monthly mean and maximum sea significant wave heights

3.4 Wavelet Power

By the use of a wavelet transform we estimate the important fluctuations in the simulations of sea level and significant wave heights. Figure 7 shows the result of the estimated wavelet power (WP) for each scale and parameter. The results from the DFA analysis are confirmed, in such way that the WP of Hs is decreasing with increasing scale, while the WP of the monthly maximum sea level is constant and the WP of the monthly mean sea level increases with scale. The difference between Hs calculated with and without sea level is weak.

The importance of the yearly cycle (1) is indicated. The second peak (2) is in the range of 3, while the third (3) has a lengths of 9 years, both maxima together represent something like the NAO cycle in the simulation. The 4th peak can only be seen in Hs but not in the sea level, the length of the scale is about 20 years. On hint to the scale can be found in Escudier *et al*. (2013). They deduced the mode of 20 years in coupled ocean-sea ice-atmosphere variability mode in the North Atlantic in a Global Climate Model. It is difficult to find a reason that the sea levels underlie no peak in this scale. The only reason one might find is a change in wind direction, further studies might go into detail.

The last pronounced peak in wavelet analysis can be observed in the sea level time series only and has a period of 30 years. It may be the response of the lateral boundary forcing of the North Sea to the multidecadal oscillation of the Atlantic Meridional Overturning Circulation (AMOC) reproduced by the Global Climate model. In a sensitive study using such kind of a model Huang *et al.* (2012) significant fluctuations were also found on that scale in the AMOC.

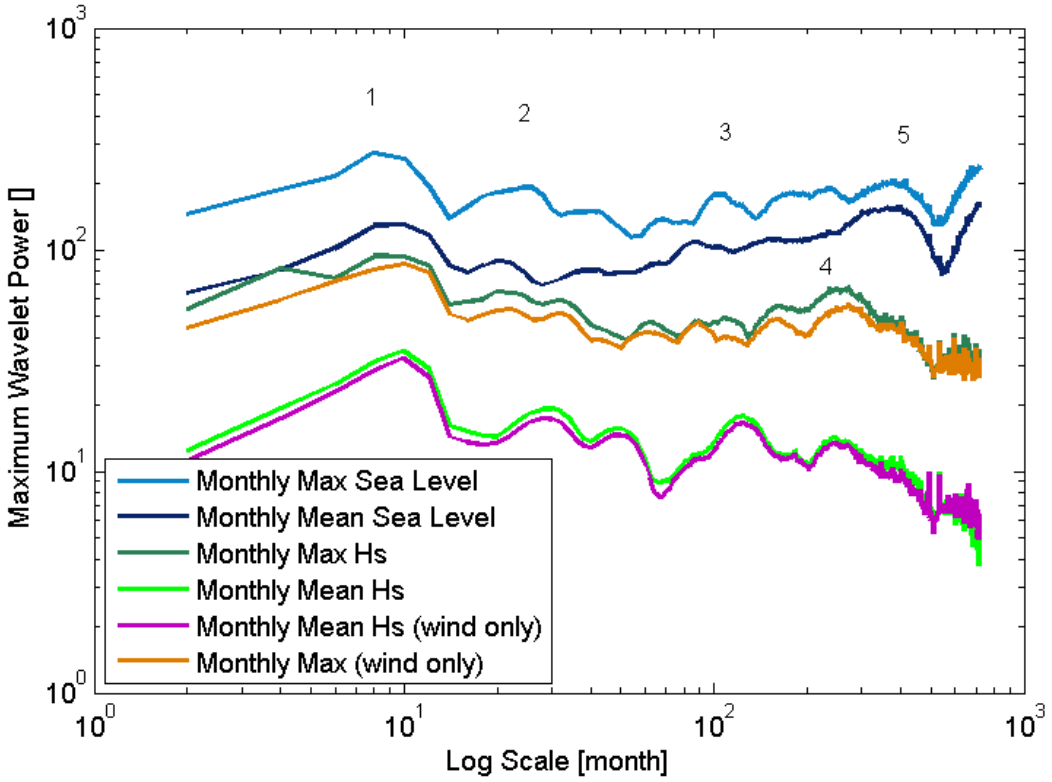


Figure 7: Wavelet Power from the monthly mean and maximum sea significant wave heights

4 CONCLUSION

Result from the so called model chain used in the KLIWAS framework are presented and analysed. Our study calculates the persistence of these downscaled climate models. The persistence found in sea level simulations is in the same order like it is known from river hydrology. The differences between the sea level simulations and the wave simulation is mainly due to long term memory which can be found in the sea level data. The typical explanation might be the slow response of the ocean to a fast atmosphere. On longer scales the persistence is not explained by atmospheric forcing. One conclusion could be drawn that stochastic analysis using the sea level – atmospheric relations must fail to reproduce the multi decadal variability.

For the sea level multi-decadal variability must be taken into account if trends in sea level are the subject of interest. For sea level maxima on longer (19 and 100 yearly) scales no trend or systematic change could be found with the technics used in this study. They are triggered by suggested random single events. For the significant wave height already the monthly mean and monthly maxima are just white noise. Therefore the long term predictability of this parameter is questionable. However, it is know that global climate models tend to underestimate the atmospheric persistence.

Process based downscaling of a global climate model into an estuary is a quite well functional method to estimate future changes of mean conditions and maxima – if an uncertainty analysis of the results is done. Non-stationary and multi-decadal hydrodynamic behaviour of an estuary to climate change are inherent uncertainties. A critical challenge in supporting adaptation is the linkage between vulnerability research and coastal management decisions with respect to this multi-scale variability.

5 ACKNOWLEDGMENTS

This work was conducted in the framework of KLIWAS www.kliwas.de. It is funded by the Federal Ministry of Transport, Building and Urban Development.

6 REFERENCES

Backhaus, J. O. (1985): *A three-dimensional model for the simulation of shelf sea dynamics*. Dt. Hydrogr. Z. 38, pp.165–187.

Bakker, A. M. and van den Hurk, B. J. (2012): *Estimation of persistence and trends in geostrophic wind speed for the assessment of wind energy yields in Northwest Europe*. Climate dynamics 39 (3-4), pp.767-782.

Berkhahn, V. and Mai, S. (2004): *Meshing Bathymetries for Numerical Wave Modelling*. Proc. of the 6th Int. Conf. on Hydroinformatics, Singapore.

Bhatthacharya, R.N., Gupta, V.K. and Waymire, E.C. (1983): *The Hurst effect under trends*, J. Appl. Probab. 20 (1983), pp.649–662.

Booij, N., R. C. Ris and L. H. Holthuijsen (1999): *A third-generation wave model for coastal regions: 1. Model descrip-tion and validation*, J. Geophys. Res. 104 (C4), pp.7649–7666.

Bunde, A., M. I. Bogachev and S. Lennartz (2012): *Precipitation and river flow: Long-term memory and predictability of extreme events, in Complexity and Extreme Events in Geoscience*, Geophys. Monogr. Ser., doi:10.1029/2011GM001112.

Dietrich, S., Winterscheid, A., Wyrwa J., Hein H, Hein, B. and Schöl A. (2013), *Development of an interdisciplinary model cluster for tidal water environments,* Geophysical Research Abstracts Vol. 15, EGU2013-5119, EGU General Assembly 2013.

Escudier, R., Mignot, J. and Swingedouw, D. (2013): *A 20-year coupled ocean-sea ice-atmosphere variability mode in the North Atlantic in an AOGCM*. Climate Dynamics 40 (3-4), pp.619-636.

Fernández, I., Hernández, C. N. and Pacheco, J. M. (2003): *Is the North Atlantic Oscillation just a pink noise?* Physica A: Statistical Mechanics and its Applications 323, pp.705-714.

Hein, H. (2008): *Vietnam Upwelling - Analysis of the upwelling and related processes in the coastal area off South Vietnam,* PhD Thesis, http://www.sub.uni-hamburg.de/opus/volltexte/2008/3931/.

Hein, H., Karfeld, B. and Pohlmann, T. (2007): *Mekong water dispersion. Measurements and consequences for the hydrodynamic modelling*, J. of Water Res. and Env. Eng., Special Issue, August 2007, pp.21-28.

Hein, H., Mai, S. and Barjenbruch, U. (2011a): Interaction of Wind Waves and Currents in the Ems Estuary. Int. Journal of Ocean and Climate Systems 2 (4).

Hein, H., Mai, S. and Barjenbruch, U. (2011b): *Coastal long term processes, tidal characteristics and climate change,* 5th International Short Conference on Applied Coastal Research, Aachen, http://www.iww.rwth-aachen.de/fileadmin/internet/scacr/SCACR_2011_Proceedings.pdf, pp.214-221.

Hein, H., Mai, S. and Barjenbruch, U. (2011c): *What Tide Gauges Reveal about the Future Sea Level*, Proc. of the 4th Conf. Acqua Alta, http://acqua-alta.de/fileadmin/design/acqua-alta/pdf/abstracts/paper/13_10/Hein_Harmut_full_papers.pdf.

Hein, H., Mai, S. and Barjenbruch, U.(2012): *Uncertainties of Drying Periods of Coarse Coastal Climate Impact Models.* Proc. of the 2nd IAHR Europe Conf., München.

Hein, H., Mayer, B., Mai, S. and Barjenbruch, U. (2013): *Process based downscaling of a global climate model into the Elbe estuary*. Proc. of 6th Int. Conf. on Water Resources and Environment Research (ICWRER), Koblenz.

Hein, H., Weiss, R., Barjenbruch, U. and Mai, S. (2010): *Uncertainties of tide gauges & the estimation of regional sea level rise*. In Proc. of the Int. Conf. Hydro.

Huang, B., Hu, Z. Z., Schneider, E. K., Wu, Z., Xue, Y. and Klinger, B. (2012): *Influences of tropical–extratropical interaction on the multidecadal AMOC variability in the NCEP climate forecast system.* Climate dynamics 39 (3-4), pp.531-555.

Hurst, H.E. (1951): *Long-term storage capacity of reservoirs.* Trans. Am. Soc. Civil Eng. 116, pp.770-799.

Lewis F.M. and Abbow, C. M. (l976): *Pyrogas from biomass. Paper presented to Conference on Capturing the Sun Through Bioconversions*, Washington, D.C.

Lye, L. M. and Lin, Y. (1994): *Long-term dependence in annual peak flows of Canadian rivers.* Journal of Hydrology 160 (1), pp.89-103.

Mai, S. (2004): *Klimafolgenanalyse und Risiko für eine Küstenzone am Beispiel der Jade-Weser-Region: Climate impact and risk assessment for the coastal zone*, Doctoral dissertation, Franzius-Inst. für Wasserbau und Küsteningenieurwesen.

Mai, S. (2008): *Statistics of Waves in the Estuaries of the Rivers Ems and Weser - Measurement vs. Numerical Wave Model*. Proc. of the 7th COPEDEC Conf., Dubai.

Mathis, M. (2013): *Projected Forecast of Hydrodynamic Conditions in the North Sea for the 21st Century*, http://ediss.sub.uni-hamburg.de/volltexte/2013/6169/.

Mudersbach, C., Wahl, T., Haigh, I. D. and Jensen, J. (2013): *Trends in high sea levels of German North Sea gauges compared to regional mean sea level changes, Continental Shelf Research*, ISSN 0278-4343, http://dx.doi.org/10.1016/j.csr.2013.06.016.

Peng, C.-K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E. and Goldberger, A.L. (1994): *Mosaic organization of DNA nucleotides*. Phys. Rev. E 49 (2), pp.1685–1689.

Pohlmann, T. (2006): *A meso-scale model of the central and southern North Sea: Consequences of an improved resolution*. Continental Shelf Res. 26, pp.2367-2385.

Sakalauskienė, G. (2003): *The Hurst phenomenon in hydrology*. Environmental research, engineering and management 3 (25), pp.16-20.

Torrence C, Compo GP. (1998): *A practical guide to wavelet analysis*. Bulletin of the American Meteorological Society 79, pp.62–78.

Wahl, T., Jensen, J., Frank, T. and Haigh, I. D. (2011): *Improved estimates of mean sea level changes in the German Bight over the last 166 years*. Ocean Dynamics 61 (5), pp.701-715.

Weisse R, von Storch, H, Niemeyer, HD and Knaack, H. (2011): *Changing North Sea storm surge climate: an increasing hazard?* Ocean Coast Manag. doi:10.1016/j.ocecoaman.2011.09.005.

Spatial Extreme Value Analysis of Significant Wave Heights Along the French Coast

Thomas Bulteau[1], Sophie Lecacheux, Alexandre Nicolae Lerma and François Paris

[1]BRGM, Orléans, France, Email: t.bulteau@brgm.fr

Abstract

Extreme value analysis is of paramount importance in coastal engineering, for structure design as well as hazard mapping. The significant wave height (SWH) is the parameter generally used to characterize the intensity of sea states. Extreme value analysis on SWH requires historical buoy records of sufficient length and good quality. However, such observation datasets are often inexistent and numerical hindcasts of waves are used instead. One advantage of using such model outputs is that an extreme value analysis over a large spatial area is possible, enabling one to highlight spatial variations on extremes.
In this study, we aim at studying spatial variations of extreme values of SWH along the French coast for current climate. We use wave data from the BoBWA-10kH database (Charles et al., 2012) which is a numerical wave hindcast for the whole West coast, performed with the third generation wave model WWIII (Tolman, 2009) and forced with ERA-40 reanalysis winds. It covers the period 1958-2002 and has a spatial resolution of 10km. Extreme value analysis is performed for about 40 points regularly distributed along the coast. The Peaks-Over-Threshold method is used and a Generalized Pareto Distribution is fitted to the data. Spatial variations along the coast for several return values of SWH (10-year, 50-year, 100-year) are presented and discussed.

1 INTRODUCTION

In coastal engineering, extreme value analysis is widely used for various applications, from flooding hazard mapping to the design of marine works. It is a way to project oneself in the future to get a sense of "what are the odds that this event could happen?" or "which event has an occurrence probability *p*?". The intensity of sea states is generally characterized by the significant wave height (SWH), which is traditionally defined as the mean wave height (trough to crest) of the highest third of the waves. Good quality long time series of SWH are required to perform a sound statistical analysis of extremes. However, the available historical buoy records along the French coast are scarce and often discontinuous, with numerous gaps occurring during extreme events like storms. To overcome this issue, numerical hindcasts of waves can be used instead. Such models are usually calibrated and validated against buoy or satellite data to provide an accurate representation of reality. Nevertheless, it is worth noting that as soon as one uses model outputs to calculate extreme values, uncertainties linked to errors and approximations (inherent to wind data and wave model simplifications) are introduced. The main advantages of using such model outputs remain (1) the length of the time series that enables one to calculate higher extreme values and to reduce confidence intervals and (2) the possibility to highlight relative spatial variations of extreme values thanks to the large spatial area covered by the data.

Numerous statistical methods exist to determine extreme wave height. Among the most commonly used, the Peaks-Over-Threshold (POT) approach has the advantage of using all the available information on extremes behavior of the time series as soon as a suitable threshold is determined. A natural candidate for the probability distribution is then the Generalized Pareto Distribution (GPD) which is the most general form of the distribution for POT samples. This method (POT-GPD) is widely used and recommended to calculate extreme values of SWH (e.g. Hawkes *et al.*, 2008; Mazas & Hamm, 2011; Li *et al.*, 2012). Nevertheless the choice of the threshold and the determination of the best law are always delicate issues, depending on the tail of the distribution. These two points must be discussed, especially when one wants to work at regional scale with a homogeneous method in order to analyze spatial variations.

Presently, ANEMOC (Numerical Atlas of Oceanic and Coastal Sea States) is the only available database of wave extreme values covering the French Atlantic and Mediterranean coasts with a good point density (Benoit *et al.*, 2006). It was realized by EDF/LNHE and the CETMEF from a wave hindcast based on the third generation model TOMAWAC (Benoit *et al.*, 1996) and the ERA-40 winds (Uppala *et al.*, 2005). It covers the period 1979-2002. A recent study of wave hindcasts intercomparison (Lecacheux & Paris, 2013) pointed out that this dataset presents a positive bias for values above the 90th percentile compared to observations. Yet, extreme value analyses are very sensible to events constituting the tail of the distribution and we can expect the ANEMOC extreme values to be overestimated.

In this study, realized for the French Ministry of Environment, we performed a spatial extreme value analysis of SWH along the French Atlantic coast using an alternative wave hindcast, namely the Bay of Biscay Wave Atlas (BoBWA-10kH) database (Charles *et al.*, 2012), and the POT-GPD method. This paper presents the preliminary results and is organized as follows: section 2 introduces the BoBWA-10kH database and the statistical analysis method; in section 3, we present the preliminary results and a comparison with the ANEMOC database; finally, section 4 is dedicated to the discussion and the conclusion. The final results of the project should be available at the end of the year.

2 DATA AND METHOD

2.1 Description and validation of BoBWA-10kH

BoBWA-10kH (Charles *et al.*, 2012) is a wave hindcast covering the period 1958-2001. It was realized with a two-way nested Wavewatch 3 (Tolman, 2009) modeling framework covering the North Atlantic (spatial resolution of 0.5°) and the French Atlantic and English Channel coasts (spatial resolution of 0.1°), and using the parameterization of Ardhuin *et al.* (2009). The model was forced by ERA-40 reanalysis winds (Uppala *et al.*, 2005) given every 6 hours at a height of 10m on a 1.125°X1.125° grid. In the simulations, the water level is supposed to be constant (mean level) and the currents are not taken into account. A calibration was carried out at the Biscay buoy on the period 1998-2002 by varying the wind input height. The results were stored hourly at the buoy locations along the coast and every six hours for each point of the grid.

The validation performed by Charles *et al.* (2012) on 9 buoys showed a good agreement with observations for the Atlantic coast but a poorer quality of the model in the English Channel. This fact was attributed to the coarse resolution of the model that prevents the proper modeling of waves coming from the North Sea and the fact that interactions with the strong tidal currents in this area was not taken into account. Lecacheux *et al.* (2013) showed that, in the area of the Bay of Biscay, BoBWA-10kH had the lowest statistical errors compared to the two other available regional hindcasts (ANEMOC and Bertin & Dodet, 2010). The highest values of wave heights (above the 90th percentile) seemed also to be better reproduced.

For this study, we investigated the capacity of BoBWA-10kH to reproduce storm events (peak, length, etc.). We compared the model outputs with observations on common periods at one offshore buoy (Biscay) and two coastal buoys (Biscarrosse and Minquiers) presented on Figure 1. For each buoy, the storm events correspond to the periods for which SWH exceeds 2/3 of the maximum value reached during the entire record. With this technic, we detected 9 events at the Biscay buoy (from 1998 to 2002), 13 events at the Biscarrosse buoy (from 1980 to 2002) and 7 events at the Minquiers buoy (from 1992 to 1994 and from 1997 to 2002). The results show a good correlation between simulations and observations ($R^2 \sim 0.87$) and we do not notice any systematic bias (cf. Figure 1). Concerning the peaks of the storms (which are of paramount importance in extreme value analysis), we noticed relative errors lower than 7 % at Biscay and Minquiers but up to 17 % at Biscarrosse. For this last buoy, the higher statistical errors can be attributed to its location close to the coast (< 5 km) and the insufficient resolution of the model in this area.

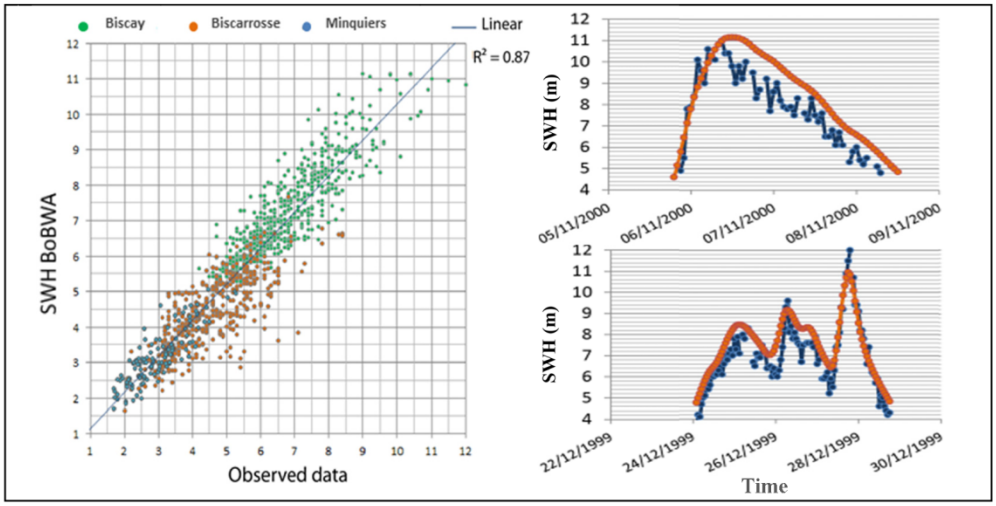


Figure 1: Comparison of BoBWA-10kH outputs with observations. Left : Correlation between measures and observations at the three buoys (Biscay, Biscarrosse and Minquiers) during the selected storm events. Right: Comparison for two storm events at the Biscay buoy

Forty three points have been selected in the BoBWA-10kH dataset to perform the statistical analysis (Figure 2): 31 grid points (6 hourly) and 12 buoys locations (hourly). They are evenly spaced of about 40-50 km along the coast and they are located about 50 km from the shore (except some coastal buoys). Nevertheless, the resolution of the model and the variations of the bathymetry along the coast did not allow selecting points at the same depth everywhere. Near the Aquitaine coast, the depth is about 50m; along the Brittany peninsula, it is around 100m; and in the English Channel it is around 30m.

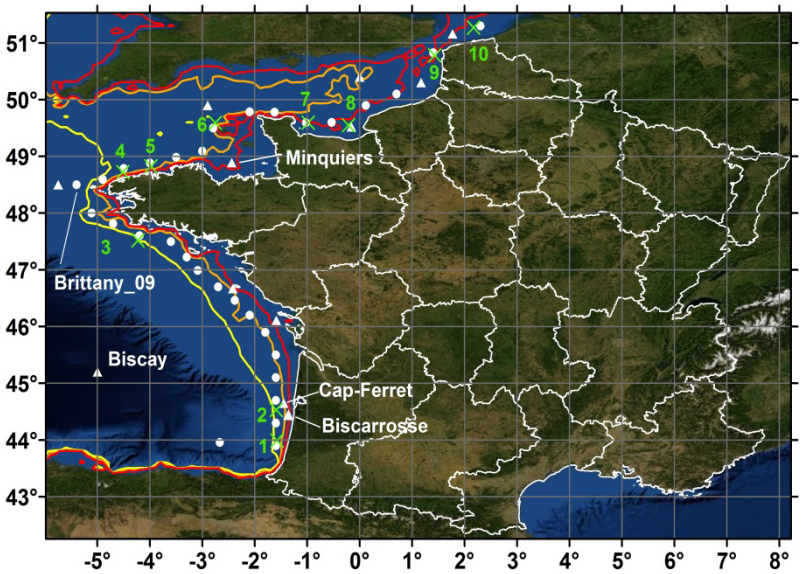


Figure 2: Selected points for the statistical analysis. White circles represent points extracted from the BoBWA-10kH grid and white triangles represent buoy locations. Green crosses correspond to the points in ANEMOC database used to compare the results of the extreme value analysis. The contours indicate the isobaths 30 m (red), 50 m (orange) and 100 m (yellow)

2.2 Method to derive the GPD

To derive extreme values from significant wave heights (SWH) time series, we first need to identify independent and identically distributed (i.i.d.) events. A simple directional analysis is used to determine whether the highest SWH are associated with several discontinuous directional sectors or not. In the first case, that means the recorded storms of a given directional sector are generated by a different kind of depressions than the ones of another sector. Therefore each storms group should be treated separately from the others as they may not be identically distributed. In practice, only a

few points in the North of the English Channel have two key directional sectors. The remaining points only have one. The independence of events is achieved using a Peaks-Over-Threshold (POT) approach combined with a temporal criterion: a minimum period of 72 hours between each peak is chosen to consider them as independent events. The POT threshold u_1 is determined roughly so as to select both weak and strong storms (which should represent a few hundreds of peaks). A statistical goodness-of-fit test (Chi²) enables us to make sure the annual occurrence of selected peaks follows a Poisson distribution at the risk level of 0.1 (when the test failed, a higher value for u_1 was chosen). A higher threshold u_2 above which storms have a statistically extreme behavior is then chosen more thoroughly using several tests and plots. We follow here the double threshold method of Mazas & Hamm (2011).

The choice has been made to fit only the Generalized Pareto Distribution (GPD) to represent the distribution of extreme wave heights along the French coast:

$$\begin{cases} P\left(SWH - u_2 > y \middle| SWH > u_2\right) = \left(1 + \frac{\xi y}{\sigma}\right)_+^{\frac{-1}{\xi}} & \text{if } \xi \neq 0 \quad (1) \\ P\left(SWH - u_2 > y \middle| SWH > u_2\right) = \exp\left(-\frac{y}{\sigma}\right) & \text{if } \xi = 0 \quad (2) \end{cases}$$

Where $s_+ = \max(s,0)$, ξ is the shape parameter and σ is the scale parameter.

It is indeed recommended to use it (Hawkes *et al.*, 2008; Li *et al.*, 2012) and since the aim of the study is a spatial analysis of extremes, it is important to be consistent and use a single statistical law for all the study sites. To determine u_2 we take advantage of the asymptotical properties of the GPD: if the sample follows a GPD then the mean excesses of SWH above u_2 vary linearly with u_2 (mean residual life plot) and the modified scale parameter $\sigma^* = \sigma - \xi u_2$ and the shape parameter ξ remain constant when u_2 increases. In practice, we fit the GPD for thresholds comprised between u_1 and a threshold corresponding to 1 event per year and search for the lowest threshold of the highest domain of linearity resp. stability (Mazas & Hamm, 2011). However, despite these theoretical properties, determining the high threshold u_2 is not always straightforward. Sometimes several values might apply. In addition to these two plots, we also perform two statistical tests namely the Kolmogorov-Smirnov and χ^2 tests for the whole range of thresholds in order to validate or invalidate the values previously determined. We focus on maximizing the p-value of the two tests and we reject the thresholds when the tests have failed (with a risk level of 0.1). Last, we draw a sensibility graph representing the variations of 100-year SWH (SWH_{100}) with respect to u_2 (Mazas & Hamm, 2011). SWH_{100} shall remain roughly constant for thresholds above u_2. During the threshold selection process, we also try to follow a recommendation of Mazas & Hamm (2011) which is to prefer a value of u_2 corresponding to about 2 events per year when the number of years is large (over 40 years), which is the case in our study. All those information enable us to determine an adequate threshold most of the time. When it still remains difficult to choose between thresholds, we also use quantile-quantile plots to make a final visual decision.

The estimation of the GPD parameters is also a crucial part of the analysis. Several methods exist, the most commonly used being the method of moments, the probability weighted moments and the maximum likelihood (Mackay *et al.*, 2011). When it comes to choosing an estimator, three characteristics need to be considered: bias (is the expected value of the estimator equal to the parameter?), efficiency (is the variance (or RMSE) of the estimator as small as possible?) and consistency (as the number of observations n increases, does the estimator value approach the true parameter value?). To stay consistent all along the French coast and to be able to do comparisons between study sites, we decided to use only one method regardless of the relative performance of the different estimators. We chose the method of moments because the corresponding estimator is the only one having a small positive bias for n around 100 and $\xi < 0$, which is typically the case for our study, while keeping the RMSE reasonably low (Mackay *et al.*, 2011). A small positive bias may compensate the slight underestimation of the GPD on grid points due to the 6-hour time step of the time series. Indeed, with a time discretization of 6 hours, the highest SWH in a given storm might be missing which can lead to an underestimation of extremes. A sensitivity test on the time step (1 hour vs 6 hours) was therefore conducted on several buoys showing a difference in the results up to 3 % on SWH_{100}.

Finally, confidence intervals are obtained from the classical Delta method (Coles, 2001). All the analyses were done under MATLAB environment with the WAFO toolbox (Brodtkorb *et al.*, 2000).

3 RESULTS

3.1 Example analysis of a study point

In this section, we go through the steps of the method to derive the GPD of SWH for the study point 'Brittany_09' located -5.4E, 48.5N (Figure 1). Figure 3 shows the waves' characteristics of the dataset. Peak directions (Dp) are represented in nautical convention (incoming direction and North = 0°). SWH values can be seen on the radial axis. The envelope of the time series data points is represented by the dashed line. The occurrence frequency of (SWH, Dp) pairs is represented by the colorbar (min = 0.02‰) with a discretization of Dp every 5° and SWH every 5 cm. It can be seen that the highest waves come from the same directional sector (≈230°-300°) so no directional analysis is required.

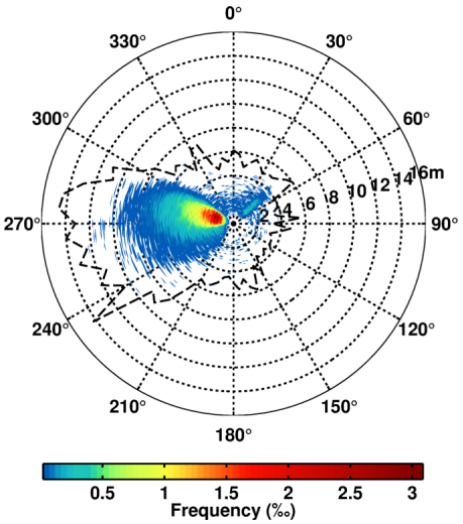


Figure 3: Polar representation of waves' characteristics for the point 'Brittany_09'

The POT threshold u_1 is fixed at 8m, after checking the resulting population of storms follows a Poisson distribution, which corresponds to 196 values. Then we have to determine the best high threshold u_2. As described in section 2.2, we adjust a GPD to the data over a wide range of thresholds (from u_1 to a threshold corresponding to 1 event per year) and look at the stability of the shape parameter ξ and of the modified scale parameter σ* (Figure 4a) and at the linearity of the mean residual life plot (Figure 4b). On Figure 4a we also display for each threshold, the corresponding number of events per year on the secondary axis.

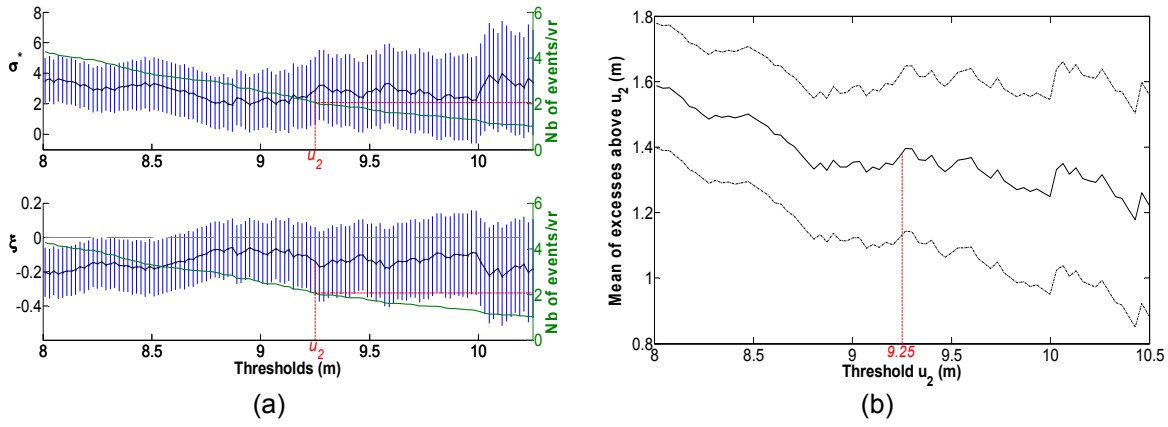
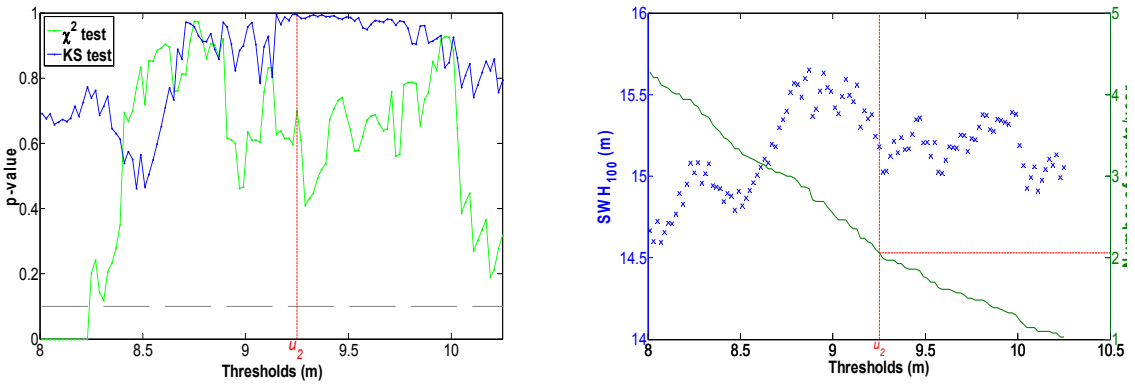


Figure 4: (a) Stability of modified scale and shape parameters for the GPD. The vertical blue bars represent the 95 % confidence intervals; (b) Linearity of the mean residual life plot. The dashed lines represent the bounds of the 95 % confidence interval

From both graphs, a value of u_2 around 9.25 m seems adequate. This value corresponds to a number of events per year of 2 or so (in accordance with the recommendation of Mazas & Hamm, 2011). One can notice a significant variation in both graphs for a value of u_2 above 10 m. However, the resulting number of events per year would be too small and so would be the corresponding number of remaining points to adjust the GPD. To validate our value of 9.25 m, we plot the variation of the p-value obtained from the χ^2 and Kolmogorov-Smirnov (KS) tests for the range of thresholds (Figure 5a). Both p-values are largely above 0.1. As a final check, we plot the variation of SWH_{100} against u_2 (Figure 5b). SWH_{100} can reasonably be considered as stable above 9.25 m.



(a) (b)

Figure 5: (a) Variation of the p-value of the two statistical tests χ^2 and Kolmogorov-Smirnov with the threshold u_2; (b) Stability of SWH_{100} with the threshold u_2

The GPD is then adjusted to the 92 data points above u_2. Figure 6 shows the result with the Hazen plotting position for the data points.

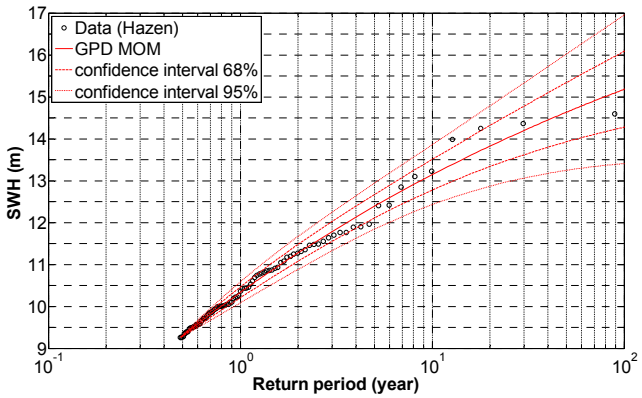


Figure 6: Generalized Pareto Distribution for the point 'Brittany_09'

3.2 Spatial analysis of extremes of SWH

Figure 7 shows the results of the statistical analysis along the French Atlantic coast (except for 3 offshore buoy locations): spatial variations of SWH_{10}, SWH_{50} and SWH_{100} are thus highlighted. We can notice that the three quantities vary the same way: the lowest values are found in the English Channel, the highest ones around the Brittany coast and values in the middle are found along the Aquitaine coast and between Brittany and Normandy. This is partly explained by the bathymetry and the varying depth at each point of the study, as described in section 2.1. Figure 7 also displays the difference between SWH_{100} and the maximum value of SWH simulated along the coast. These two quantities vary similarly along the coast with a range of differences from about 0.6 m (in deep and relatively exposed areas such as the West of Brittany) to about 0 m (in shallow and not very exposed areas such as the North of the English Channel or the Normandy coast to the East of Cotentin, but also in deep and more exposed area such as the South of Brittany). A difference close to 0m suggests that historical events have generated waves with SWH close to the 100-year value.

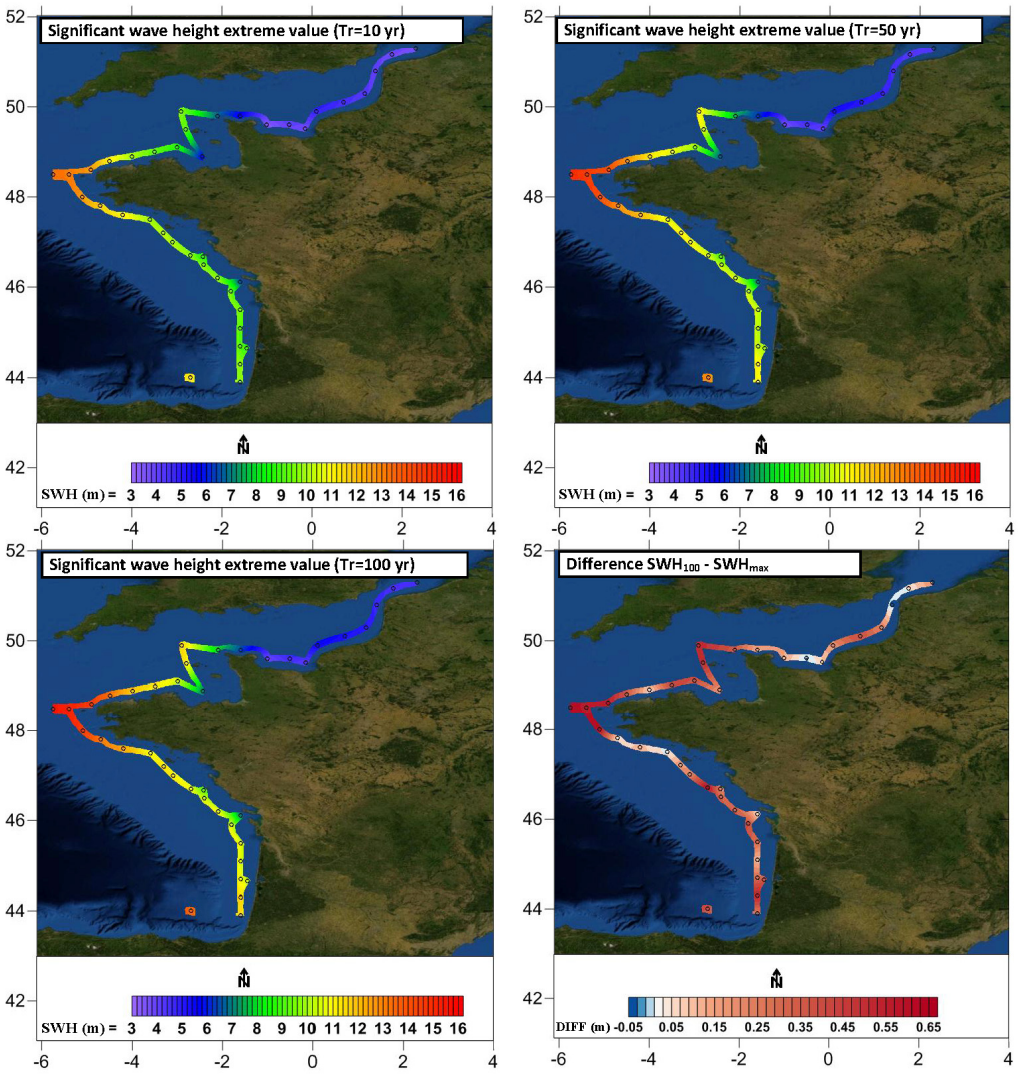


Figure 7: Results of the statistical analysis for 40 points along the coast: (1) Values of SWH for return periods of 10 years (up left), 50 years (up right), and 100 years (down left) (2) Differences between SWH_{100} and SWH_{max}

3.3 Comparison with ANEMOC database

The spatial extreme value analysis performed with the BoBWA-10kH dataset is then compared with the ANEMOC product. As mentioned in the introduction, it is likely that results of extreme value of SWH obtained with the ANEMOC database will be higher than those obtained with BoBWA-10kH. Indeed, Lecacheux & Paris (2013) showed that ANEMOC dataset presents a positive bias for values above the 90th percentile and that BoBWA-10kH compared better with observations for this range of values.

The extreme value analysis in ANEMOC is similar to the one presented in this paper. Storms were selected with a POT approach and two distributions were adjusted to the data: a GPD (with the maximum likelihood estimators) and the exponential distribution. To be able to compare results, we focus only on the GPD. However, since the meshes that were used in the models are different, both in nature and resolution, it was not possible to perform a comparison of perfectly co-localized points. In addition, the extreme analysis in ANEMOC was performed only for a selection of points. Nevertheless, we managed to do the comparison for 10 points along the coast (cf. green crosses on Figure 1 that represent ANEMOC points). Results are presented in Figure 8.

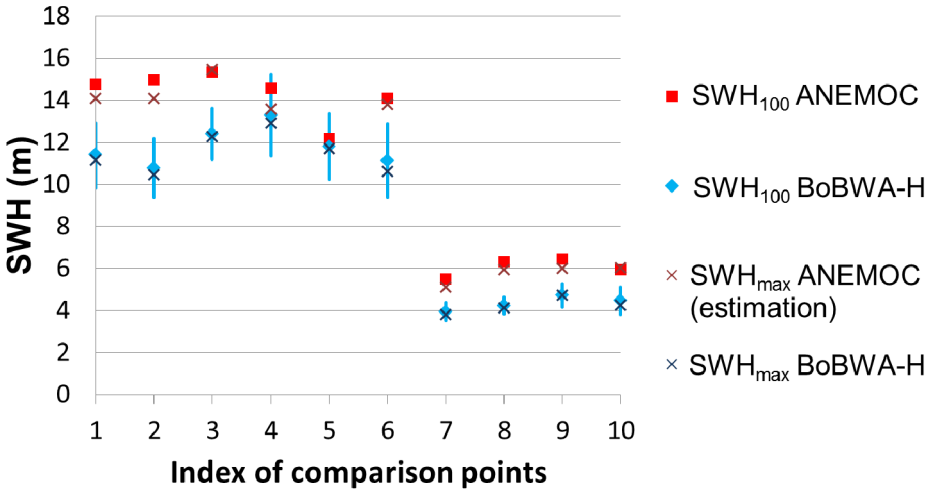


Figure 8: Comparison of extreme values obtained with ANEMOC and BoBWA-10kH. The indices of comparison points correspond to the green crosses on Figure 1. Vertical blue bars represent 95% confidence intervals on SWH_{100} obtained with BoBWA-10kH

We first noticed the same regional tendencies between BoBWA-10kH and ANEMOC because the values vary similarly. Second, and as foreseen, values of SWH_{100} of ANEMOC are higher than those obtained with BoBWA-10kH. The difference is up to 3-4 m along the Aquitaine and Brittany coasts (points 1 to 6, between 50 m and 100 m depth) and around 2 m in the English Channel (points 7 to 10, 30 m depth). Even if the comparison is partial because of the above described constraints (no co-localization, influence of the depth on the results) that limit the number of comparison points, it still highlights a significant difference of about 2m in average all along the French Atlantic coast between the values of SWH_{100} obtained with the two hindcasts. As it can be seen in Figure 8, ANEMOC values of SWH_{100} are much higher than the upper bound of the calculated confidence intervals (except for points 4 and 5). Therefore, the difference cannot be explained with the help of probabilities. In addition, we also see that SWH_{max} follows almost the same tendency than SWH_{100} whatever the database. Thus, it appears that the initial raw data (model outputs) is of paramount importance for extreme value analysis and probably more important than practical questions related to statistics and probability theory (which distribution?, Which threshold?, etc.).

4 DISCUSSION AND CONCLUSION

The objective of this study was to perform a spatial extreme value analysis of SWH along the French Atlantic coast, taking advantage of a recent numerical wave hindcast BoBWA-10kH. That would offer an alternative to ANEMOC which is currently the only available database of wave extreme values covering the French Atlantic and Mediterranean coasts with a good point density (Benoit *et al.*, 2006).

By comparing extreme values obtained with both datasets along the coast, we found the same general spatial pattern but BoBWA-10kH values were in average 2 m below ANEMOC values. This result underlines the crucial role of models and their calibration and validation to accurately represent storm peaks in order to have good quality data and to perform a sound extreme value analysis.

Concerning the method to derive the GPD, we decided to use exclusively the method of moments to estimate the GPD parameters because it allows the comparison between study sites and it may compensate the slight underestimation of extremes due to the 6-hour time step for grid points (see section 2.2). If the goal was not to produce a regional map of extreme values but to derive extremes on a specific site, another method (such as maximum likelihood or probability weighted moments) might be more appropriate and we should perform a proper comparison of the results to choose the best method.

It is worth noting that the extreme value analysis is based on 44 years, from 1958 to 2001. Thus,

several recent impacting storm events, such as Klaus in 2009 or Xynthia in 2010, are not taken into account in the analysis. For example, the Cap Ferret buoy (see Figure 1) recorded a value of SWH as high as 11.3 m during Klaus in January 2009 (CETMEF, 2012). This value is about 90 cm above the calculated SWH_{100} in our study. It is still below the upper bound of the 95 % confidence interval (11.8m), but there is no doubt that if the time series had included 10 more years, the resulting extreme values would have been different. This observation raises another issue: since long numerical hindcasts such as BoBWA-10kH or ANEMOC are not regularly updated and completed with more data, how can we include all the available information on SWH in the extreme value analysis? How to combine time series with specific information (qualitative or quantitative information on historical events)? In the related field of hydrology, probabilistic methods have been developed over the last 30 years for the consideration of historical floods and revealed the real added value that represents the historical information, incomplete as it may be (Gaume *et al.*, 2010; Payrastre *et al.*, 2012; N'Guyen *et al.*, 2013). The application of these methods in the field of coastal risks and more particularly in the study of extreme wave heights still falls within the area of research and should be investigated in future works to improve extreme value analyses.

5 ACKNOWLEDGMENT

This work is funded by the French Ministry of Environment. The BoBWA-10kH database was realized within a research project hold by the BRGM and CNRM-GAME (Météo-France-CNRS) and the PhD of Elodie Charles funded by the AXA Research Fund.

6 REFERENCES

Ardhuin, F., Chapron, B. and Collard, F. (2009): *Strong decay of steep swells observed across oceans*, Geophysical Research Letters 36 (6), art. n°. L06607. doi:10.1029/2008GL037030.

Benoit, M., Lafon, F. and Goasguen, G. (2006): *Constitution et exploitation d'une base de données d'états de mer le long des côtes Françaises par simulation numérique sur 23 ans,* in *Actes des IXèmes Journées Nationales Génie Côtier Génie Civil*, Brest (France), September 12-14, 2006, pp.21-30. doi:10.5150/jngcgc.2006.003-B.

Benoit, M., Marcos, F. and Becq, F. (1996): *Development of a third generation shallow water wave model with unstructured spatial meshing*, Proceedings of the 25th International Conference on Coastal Engineering (ICCE'1996), Orlando (Florida, USA), September 2-6, pp.465-478.

Bertin, X. and Dodet, G. (2010): *Variabilité du climat de houle dans le Golfe de Gascogne au cours des six dernières décennies*, in *Actes des XIèmes Journées Nationales Génie Côtier Génie Civil*, Les Sables d'Olonne (France), June 22-25, 2010. doi:10.5150/jngcgc.2012.005-B.

Brodtkorb, P.A., Johannesson, P., Lindgren, G., Rychlik, I., Rydén, J. and Sjö, E. (2000): *WAFO - a Matlab toolbox for analysis of random waves and loads*, Proceedings of the 10th International Offshore and Polar Engineering conference, Seattle, Vol III, pp.343-350.

CETMEF (2012): *Fiches synthetiques de mesure des états de mer*, Centre d'Etudes Techniques Maritimes et Fluviales, February 2012.

Charles, E., Idier, D., Thiébot, J., Le Cozannet, G., Pedreros, R., Ardhuin, F. and Planton, S. (2012): *Present Wave Climate in the Bay of Biscay: Spatiotemporal Variability and Trends from 1958 to 2001*, Journal of Climate 25 (6), pp.2020-2039. doi:10.1175/JCLI-D-11-00086.1.

Coles, S. (2001): *An Introduction of Statistical Modelling of Extreme Values*, Springer, London.

Gaume, E., Gaal, L., Viglione, A., Szolgay, J., Kohnova, S. and Blöschl, G. (2010): *Bayesian MCMC approach to regional flood frequency analyses involving extraordinary flood events at ungauged sites*, Journal of Hydrology 394, pp.101-117. doi:10.1016/j.jhydrol.2010.01.008.

Hawkes, P.J., Gonzales-Marco, D., Sánchez-Arcilla, A. and Prinos, Panayotis (2008): *Best practice for the estimation of extremes: A review*, Journal of Hydraulic Research, 46 (2), pp.324-332.

Lecacheux, S. and Paris, F. (2013): *Projet Climats de houle phase 1 : Intercomparaison de simulations rétrospectives et prospectives dans le Golfe de Gascogne*, Rapport BRGM/RP-61651-FR, 116 p., 58 ill, 1 ann.

Li, F., Bicknell, C., Lowry, R. and Li, Y. (2012): *A comparison of extreme wave analysis methods with 1994-2010 offshore Perth dataset*, Coastal Engineering 69, pp.1-11. doi:10.1016/j.coastaleng.2012.05.006.

Mackay, E.B.L., Challenor, P.G. and Bahaj, A.S., (2011): *A comparison of estimators for the generalised Pareto distribution*, Ocean Engineering 38, pp.1338-1346.

Mazas, F. and Hamm, L. (2011): *A multi-distribution approach to POT methods for determining extreme wave heights*, Coastal Engineering 58, pp.385-394.

N'Guyen, C.C., Payrastre, O. and Gaume, E. (2013): *Inventaires de crues extrêmes sur des sites non jaugés et analyse statistique régionale des débits: réflexions méthodologiques et évaluation des performances*, La Houille Blanche, n°2, pp.16-23.

Payrastre, O., Gaume, E. and Andrieu, H. (2012): *Information historique et étude statistique des crues extrêmes : quelles caractéristiques souhaitables pour les inventaires de crues historiques ?* in *Actes du Congrès SHF : « Evènements extrêmes fluviaux et maritimes »*, Paris, February 1-2, 2012.

Tolman, H.L., (2009): *User manual and system documentation of WAVEWATCH III version 3.14*, NOAA/NWS/NCEP/MMAB Tech. Note 276.

Uppala, S. M. et al. (2005): *The ERA-40 re-analysis*. Quart. J. Roy. Meteor. Soc. 131, pp. 2961–3012. doi:10.1256/qj.04.176.

Copula functions as a useful tool for coastal engineers

Thomas Wahl[1,2], Jens Bender[1,3] and Jürgen Jensen[1,3]

[1]Institute of Advanced Studies – FoKoS, University of Siegen, Germany, Email: thomas.wahl@uni-siegen.de

[2]College of Marine Science, University of South Florida, USA

[3]Research Institute for Water and Environment, University of Siegen, Germany

Abstract

This paper discusses how multivariate statistical models can be useful for coastal engineers, by providing a brief introduction and outlining various obvious applications of Copula models. Especially when assessing the hydraulic loading factors for coastal and offshore infrastructure the joint occurrence of different phenomena (e.g. storm surge water level and waves) may be responsible for the failure of the structure, and should therefore be thoroughly examined during the design process. This and various other potential applications are discussed in the paper. We also present an example where different bivariate and a trivariate Copulas are used to analyze a multivariate dataset consisting of storm surge water levels, significant wave heights, and wave peak periods measured offshore Sylt Island, the biggest German North Sea island located in the northeast of the German Bight. The bivariate Copulas are used to determine the joint exceedance probabilities for the different pairs of the variables of interest and a trivariate Copula is constructed to jointly analyze all three parameters.

1 INTRODUCTION

Extreme value analyses (EVA) play an important role in the daily work of many coastal engineers, as they are usually part of the design process of coastal (and offshore) infrastructure. In the past, coastal structures were designed using simple deterministic approaches, whereas nowadays, design values are usually derived from statistical analyses. Some countries have already implemented probabilistic and risk based design methods, which also requires the estimation of the relevant hydraulic loading parameters and the associated exceedance probabilities. The latter are often derived by fitting parametric univariate statistical models to some sort of observational (or simulated) data set, e.g. water level measurements from tide gauges if the storm surge water level is the parameter of interest. This one-dimensional procedure to derive the exceedance probability ignores that there are usually multiple parameters which can occur simultaneously and have a potential effect on the reliability of coastal structures (e.g. extreme total still water levels coincide with high wind waves). Hence, in order to derive reliable exceedance probabilities for the design process, it is preferred that all parameters being relevant for the particular investigation area and application are jointly analyzed in a multivariate statistical framework.

Dealing with multivariate problems generally raises two important issues: (1) *the dependence* of the marginal parameters, and (2) *their distributions*. Figure 1 shows that the characteristics of the marginal parameters prescribe the procedure which has to be chosen for calculating the joint exceedance probabilities. In terms of the *dependence*, two situations are possible; the considered parameters can be statistically independent or dependent (Salvadori *et al.*, 2007). Chi-squared tests (or other mathematical tests) can be used to test the null-hypothesis of independency (e.g. Greenwood & Nikulin, 1996). If the test results support the assumption, joint probabilities can be simply calculated by multiplying the exceedance probabilities of the marginal parameters derived from univariate EVA. Due to the simplicity of this approach, independency between parameters is sometimes falsely assumed when it comes to practical applications. Furthermore, various approaches to transform marginal variables in a way that they can be treated as independent are described in the literature (e.g. Dixon & Tawn, 1995), but not further discussed here. If the assumption of independency does not hold (e.g. the correlation r is significantly different from zero) one has to choose a multivariate model to calculate the joint probabilities.

The ***distributions*** of the marginal parameters can be from the same or from different families. If the variables under investigation follow a (univariate) distribution from the same family (e.g. all parameters are Gumbel or Normally distributed), a corresponding multivariate statistical models (e.g. bivariate Gumbel or bivariate Normal model) can be used. The literature also describes methods to transform marginal variables in a way to achieve distributions belonging to the same family (e.g. Hutchinson & Lai, 1991), but these are not further discussed here.

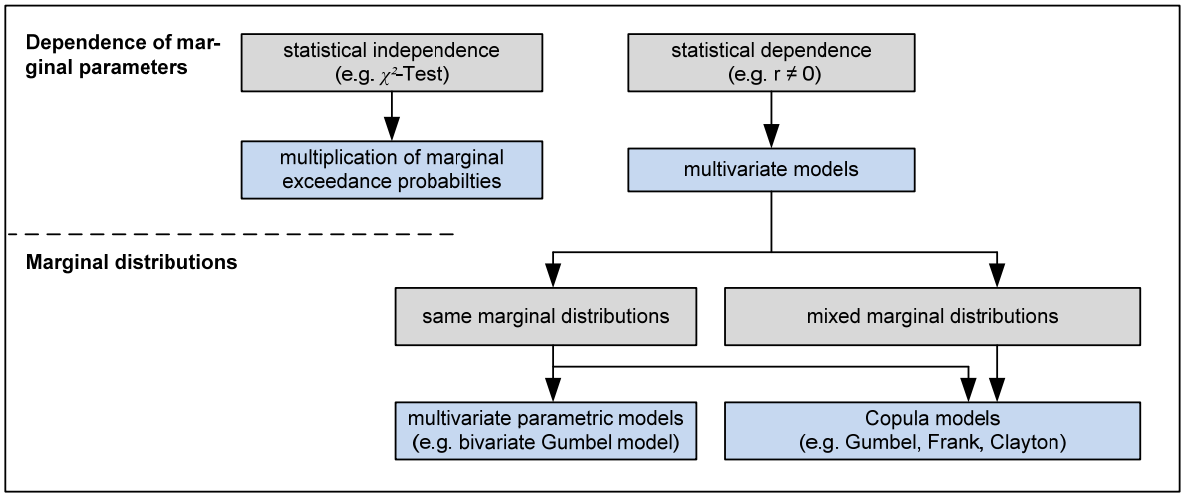


Figure 1: Effects of the characteristics of the marginal parameters on joint probability analyses

In reality, when dealing hydro-meteorological data sets, for example when assessing the hydraulic loads for flood defense structures, the variables of interest are usually not independent from each other and often differently distributed. In these cases Copula functions can be used as an alternative to traditional multivariate parametric models. Copula functions allow the joint analysis of dependent variables with mixed marginal distributions and are therefore very flexible and widely applicable. This paper outlines how Copulas can be useful for various coastal engineering purposes, a discipline where multivariate problems often arise and where Copulas are not yet commonly used. The latter will be shown in Sect. 3, after providing a very short and basic introduction to the Copula concept in the following Sect. 2. In Sect. 4, we will present an example where bivariate and trivariate Copula models are used to calculate joint exceedance probabilities for a set of design relevant variables for many coastal structures. Finally, a brief summary and some conclusions will be given in Sect. 5.

2 THEORETICAL BACKGROUND AND THE ARCHIMEDEAN COPULA FAMILY

Sklar (1959) describes the connection between a Copula C and a bivariate cumulative distribution function (cdf) $F_{XY}(x,y)$ of any pair (X,Y) as follows (note that the extension to a more general d-dimensional framework is possible):

$$F_{XY}(x,y) = C[F_X(x), F_Y(Y)] \tag{1}$$

where $F_X(x)$ and $F_Y(y)$ are the univariate marginal distributions. Further important features of Copulas and information on the theoretical background can be found for example in Nelsen (1999), who provides a detailed introduction to the subject.
A large number of different Copula functions exist, which in turn belong to different Copula families, e.g. the Elliptical, the Normal, the t-Student, or the Archimedean family. Copulas belonging to the latter are most often used in hydrological science (e.g. Favre *et al.*, 2004) as they are flexible, easy to construct, and able to model a wide range of dependence structures. Therefore, we also use Archimedean Copulas for the example given in Sect. 4; in particular we consider the Clayton (lower tail dependence), Frank (no tail dependence), and Gumbel (upper tail dependence) Copulas.

Archimedean Copulas are constructed with the Copula generators φ: [0,1] → [0,∞], which are strictly monotonically decreasing functions with φ(1) = 0 (e.g. Nelsen, 1999):

$$C_\theta(u,v) = \varphi^{-1}\left[\varphi(u)+\varphi(v)\right] \text{ , with } u = f_x(x);\ v = f_y(y) \tag{2}$$

Table 1 summarizes the relevant information on the Copula functions used for the present study. The generator functions φ(t) are displayed, as well the relationships between the Copula parameters θ and the rank correlation Kendall's τ. The latter represents a well-known nonparametric measure of dependence and can be calculated from the available observations. With this information the Copula parameters θ can be calculated.

Table 3: Archimedean Copula functions considered for the present study and their generator functions, ranges for the Copula parameters θ, and connections to Kendall's τ

Copula function C_θ	Generator φ(t)**	Range θ ∈	Kendall's τ
Clayton (or Cook-Johnson)			
$\left[u^{-\theta}+v^{-\theta}-1\right]^{\frac{-1}{\theta}}$	$t^{-\theta}-1$	$[0,\infty)$	$\frac{\theta}{\theta+2}$
Frank			
$-\frac{1}{\theta}\ln\left[1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{e^{-\theta}-1}\right]$	$-\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$	$(-\infty,\infty)\setminus\{0\}$	$1-\frac{4}{\theta}\left[1-D_1(\theta)\right]^*$
Gumbel (or Gumbel-Hougaard)			
$\exp\left\{-\left[(-\ln u)^\theta+(-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\}$	$(-\ln t)^\theta$	$[1,\infty)$	$1-\theta^{-1}$

* 1. Debye Function $D_1(\theta)=\frac{1}{\theta}\int_0^\theta \frac{t}{e^t-1}dt$

** $t = u$ or $t = v$

3 COPULAS IN COASTAL ENGINEERING

Although first mentioned more than 50 years ago, it was not before the 1990s and early 2000s that Copulas experienced a breakthrough in various scientific disciplines, including financial risk management, econometrics, and hydrology to name a few. Especially in the early years of this century the number of papers dealing with Copulas and topics related to water science (e.g. analyses of rainfall, river floods, droughts etc.) has significantly increased. An interesting website providing a (nearly) complete compilation of these papers is hosted by the Statistics in Hydrology (STAHY) group's website (www.stahy.org). Fig. 2 was compiled based on the available information from this website and highlights that Copulas made their way to the hydrology sector throughout the last decade (for the year 2013 papers that were published or in press by August 2013 were included). On the other hand, the number of Copula papers dealing with coastal engineering topics has been and is still very limited (some of the coastal engineering papers were found through an additional literature study and added to those already referenced on the STAHY website).

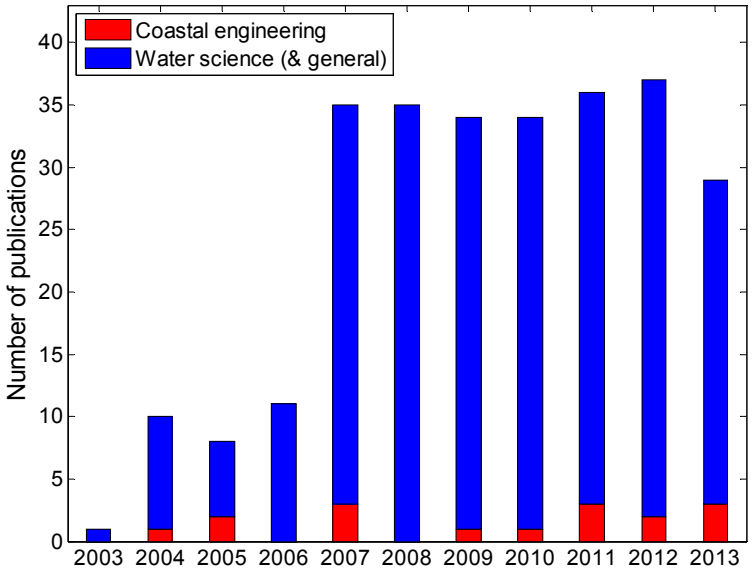


Figure 2: Number of peer-reviewed publications dealing with Copulas and being related to water science (some of the papers are more general) and coastal engineering

This is very surprising given that multivariate problems are quite common for coastal engineers and regularly arise for example during the design process of coastal and offshore infrastructure or when planning beach nourishment projects to cope with coastal erosion. Some obvious examples where Copulas can be potentially useful for coastal engineers are summarized in Fig. 3 and briefly discussed in the following.

A typical problem when calculating return periods of extreme still water levels consists in the separation and joint analysis of the deterministic astronomical tidal signal and the stochastic surge component. These two variables often show a certain dependence structure that needs to be accounted for during the statistical analyses. In the past, this has been achieved by transforming the data in a way to derive independency, which then allowed the simple multiplication of the (univariate) exceedance probabilities of the two marginal parameters (the so-called Revised Joint Probability Method, see Tawn and Vassie, 1989). More recently, an alternative method has been introduced and is now widely applied in the United Kingdom (UK), which uses the skew surge (i.e. the difference between the observed water level and the closest astronomical high water) instead of the actual surge. This also assures independency between the variables under investigation. The obvious approach of using Copulas to jointly analyze the astronomic tidal water levels and surges (despite their dependency) has yet to be tested. When the coastal flood risk is assessed, this usually includes the calculation of failure probabilities of existing defense structures and the determination of the damages (or adverse consequences) in the hinterland. Both the failure probabilities and damages are functions of the height of a storm surge, and its duration or intensity (e.g. Salecker *et al.*, 2011; Wahl *et al.*, 2010, 2011, 2012a, 2012b), which again leads to a multivariate problem. In estuarine areas and bays the water level observed during a storm surge event is often influenced by the river discharge (see e.g. Zhong *et al.*, 2013) and/or the amount of rain which enters the water body directly or with a time-lag as surface run-off. When analyzing beach erosion rates (or the reliability of dunes), the inter-arrival time between successive storms also plays an important role as it determines to what extent the beach can recover between events. Furthermore, the sea state often dominates the erosion risk and therefore it is preferable to jointly analyze storm surge and wave parameters (e.g. the significant wave height, wave period, and wave direction) (Corbella & Stretch, 2012). These are all examples where multivariate problems arise while assessing the hydraulic loading for coastal man-made and natural structures. Besides this, the calculation of failure probabilities also requires the 'combination' of probabilities of the exceedance of different limit state functions or the occurrence of different failure mechanisms (e.g. Kortenhaus *et al.*, 2002). The current state-of-the-art for calculating failure probabilities does not take potential dependencies between failure mechanisms into account, they are either assumed to be independent or perfectly dependent (i.e. if one failure mechanism occurs, another one automatically also occurs), whereas in reality there might be some kind of weak, moderate, or strong dependency that could be modeled through Copula functions.

These examples outline that calculating joint probabilities is required for various standard coastal engineering applications. The decision which parameters are relevant in a particular case and should therefore be considered within the statistical assessment resides with the practitioner/researcher.

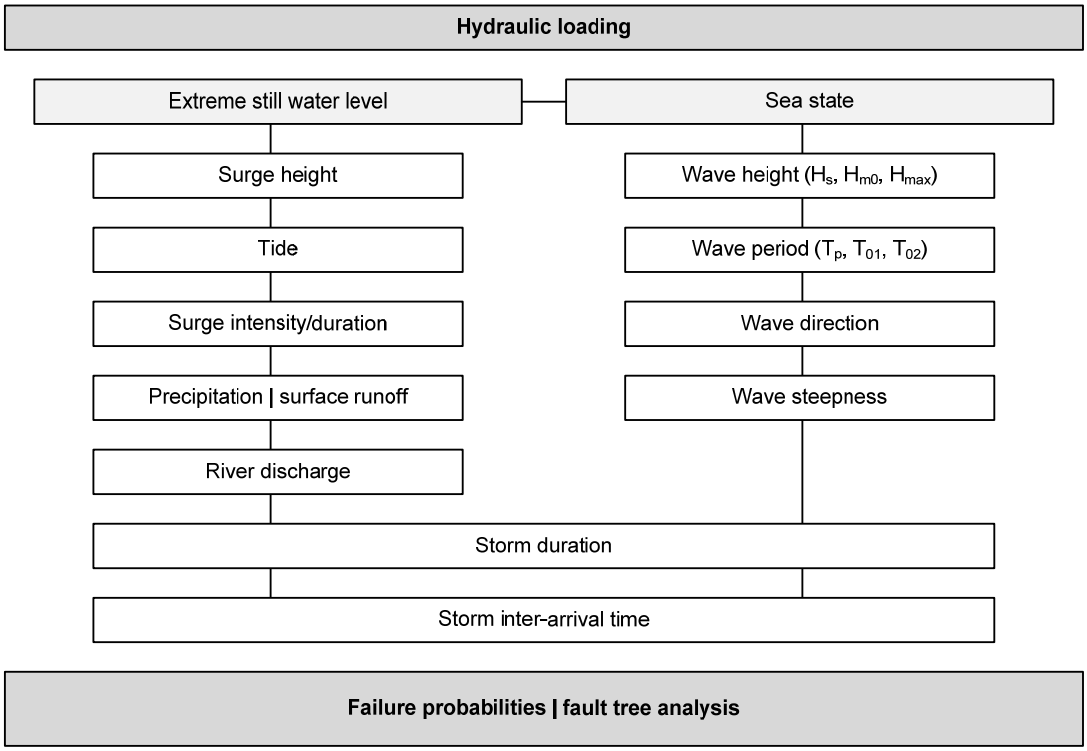


Figure 3: Examples for possible applications of Copula functions in coastal engineering

4 EXAMPLE

In this section we present an example using different Archimedean Copulas (see Tab. 1) and a fully nested trivariate Copula to analyze a data set consisting of simultaneous water level and wave measurements taken on the west side of Sylt Island (offshore Westerland) in the north-eastern German Bight between 1993 and 2008. The parameters of interest are the water level heights (we are only interested in extreme events), the significant wave heights and peak periods. These are three important input parameters for dike and/or dune breach models and they determine the failure probability of a structure and potential damages in the hinterland. The dominating variable is the storm surge water level. Therefore we search for tidal high waters where the water level exceeded a value of 1.5 m above Normal Null (NN; the German Ordnance datum which lies nowadays ~15cm above mean sea level) and the significant wave heights and peak periods observed closest to the high water time. The water threshold was derived from residual life and parameter stability plots (not shown here) as described by Coles (2001). In order to assure independency between events, two tidal high waters had to be below the threshold between adjacent storm surges. This event-selection procedure results in a data set consisting of 95 events for the analyses.

4.1 Bivariate analysis

First, different univariate distribution functions are fitted to the storm surge water levels (S), peak periods (Tp), and significant wave heights (Hs). From goodness of fit tests (GoF; not further discussed here) the Generalized Pareto distribution (GPD) was selected for S, the Normal distribution for Tp, and the Generalized Extreme Value distribution (GEV) for Hs. The empirical exceedance probabilities (here, Gringorten's plotting positions (Gringorten, 1963)) and theoretical distribution functions are shown in Fig. 4. The distribution parameters were estimated with the Maximum Likelihood method.

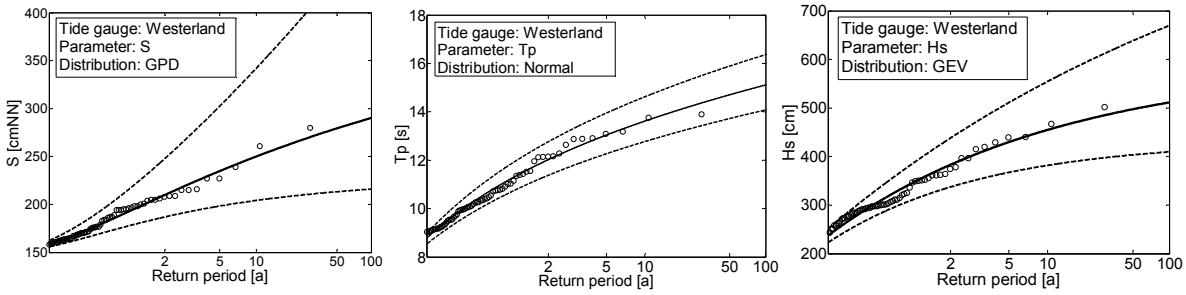


Figure 4: Marginal distributions for the parameters peak water level S (left), peak period Tp (middle) and significant wave height Hs (right)

Knowing the marginal distributions, it is now tested if the variables of interested can be assumed to be independent from each other. Therefore we use different measures of dependence, namely the Pearson correlation coefficient r and the rank correlation coefficients Spearman's ρ and Kendall's τ. The correlation coefficients and associated p-values are shown in Tab. 2 and it is obvious that all three parameters are significantly correlated with each other. The strongest correlation exists between Hs and Tp, whereas the other two pairs show a similar correlation (with the one for S and Tp being slightly higher than for S and Hs).

Table 2: Different correlation measures and p-values (in brackets) between the parameter pairs S|Hs, S|Tp, and Hs|Tp

Pair	**r**	ρ	**τ**
S | Hs	0.56 ($4 \cdot 10^{-9}$)	0.51 ($2 \cdot 10^{-7}$)	0.35 ($7 \cdot 10^{-7}$)
S | Tp	0.58 ($5 \cdot 10^{-10}$)	0.53 ($5 \cdot 10^{-8}$)	0.37 ($2 \cdot 10^{-7}$)
Hs | Tp	0.70 ($2 \cdot 10^{-15}$)	0.70 ($6 \cdot 10^{-14}$)	0.50 ($1 \cdot 10^{-12}$)

The scatter plots in Fig. 5 furthermore highlight that the pairs S | Hs and S | Tp have tail dependency in the upper right, while the third pair Hs | Tp has no evident tail dependency. From the three Archimedean Copulas considered for the present study (see Tab. 1) the Gumbel Copula has tail dependency in the upper right and the Frank Copula has no tail dependency. Hence it can be assumed that these Copulas may be appropriate to calculate the joint exceedance probabilities for the different variable pairs. This assumption is supported by the fact that the grey dots in Fig. 5, which were simulated based on the Gumbel (for S | Hs and S | Tp) and Frank (for Hs | Tp) Copulas, are in reasonable accordance with the observed data pairs (black dots in Fig. 5). This simple graphical GoF test has been used in earlier studies (Serinaldi & Grimaldi, 2007; Klein *et al.*, 2008; Wahl *et al.*, 2012), but the interpretation of the results is complicated when the observational data set is relatively small. There are, however, nowadays many other and much more sophisticated GoF tests available and described in various papers (e.g. Berg, 2009; Genest *et al.*, 2009). This important topic of selecting appropriate Copula functions by applying and evaluating different GoF tests is beyond the scope of this present paper. Thus, we use the Gumbel Copula to calculate the joint exceedance probabilities for the pairs S | Hs and S | Tp, and the Frank Copula for the pair Hs | Tp. The results are shown in Fig. 5 as isolines for selected (and design relevant) exceedance probability levels. The Copula parameters derived for the different variable pairs are listed in the lower right panel of Fig. 5. The isolines displayed in the figure represent the probability that both variables exceed a given threshold at the same time. This 'AND' case is more relevant for the design of most coastal defense structures (e.g. dikes) than the alternative 'OR' case, where the probability is assessed that only one of the variables exceeds a certain threshold. Anapproach of calculating return periods from multivariate data sets using the Kendall function (and associated Kendall return periods) was presented by Salvadori *et al.* (2011) and most recently introduced to the coastal engineering community (Salvadori *et al.*, 2013), providing a consistent multivariate framework for designing coastal/offshore infrastructure or performing safety checks/risk assessments for existing structures. This may encourage practitioners and scientists to use multivariate statistical models for a wider range of coastal engineering applications.

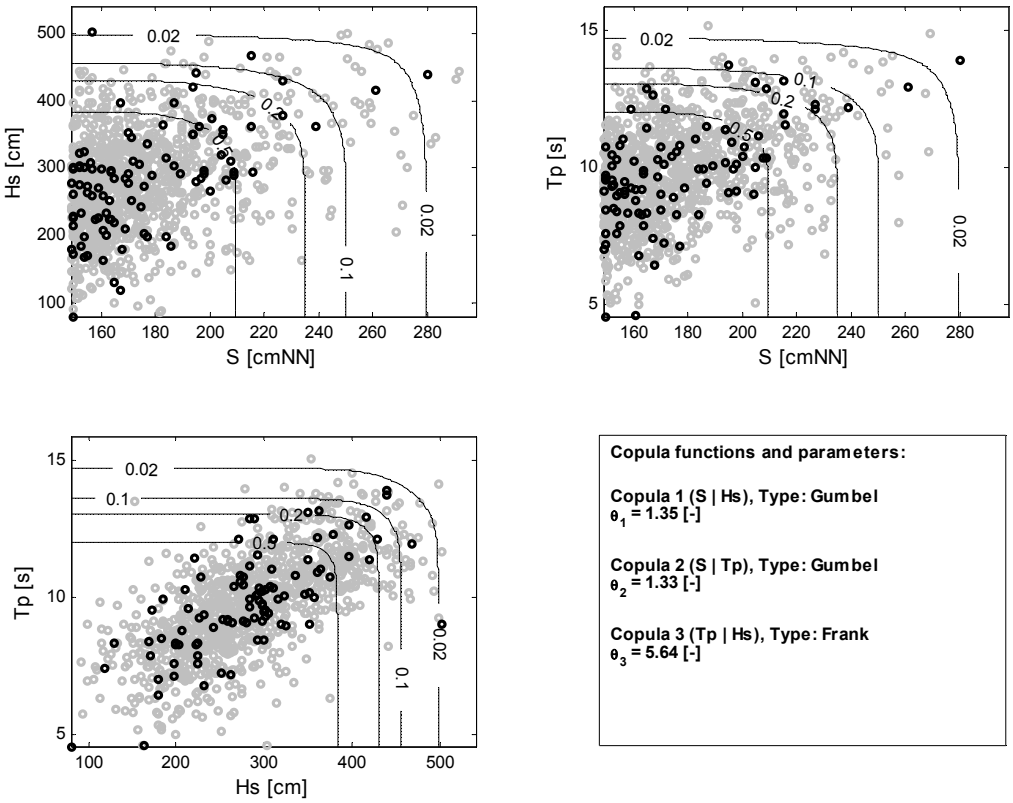


Figure 5: Results from Copula simulation (grey dots) for the pairs (S|Hs), (S|Tp) and (Hs|Tp) and contours from bivariate statistical analyses

4.2 Trivariate analysis

It has been mentioned above that the extension of Eq. (1) to the d-dimensional case is straightforward. Therefore, it is possible to consider three (or more) relevant variables within the statistical assessment and in the following we provide an example for the trivariate analysis of all three parameters described in the previous section. We use a fully nested approach to construct the trivariate Copula (see for example Serinaldi & Grimaldi, 2007; Corbella & Stretch, 2012; Wahl *et al.*, 2012). In this approach one dimension is added step by step, first a 2-Copula is derived for two of the variables, and then the remaining variable is added through a second Copula. This approach is restricted in its flexibility, given that there is always one Copula less then variables analyzed. In our case there are three variables, but only two mutual bivariate dependence structures can be freely specified through the two Copulas, i.e. the last variable included in the model shares the same Copula with the other two variables and therefore should also have a similar dependence structure (or rank correlation) with both of them. For the purpose of the present study we shall assume that this criteria is fulfilled given that the dependence measures for the variable pairs (S | Hs) and (S | Tp) are quite similar. Furthermore, the fully nested approach requires that the degree of dependence decreases with the level of nesting, and this is also fulfilled by the data set under consideration, as the correlation of the pair Tp | Hs is larger than for the other pairs.
Hence, we use the following fully nested trivariate Copula that is constructed from a Frank Copula (for the pair Tp | Hs) and a Gumbel Copula (to include S):

$$C(u_1,u_2,u_3) = \varphi_2^{-1}\left(\varphi_2\left(\varphi_1^{-1}[\varphi_1(u_1)+\varphi_1(u_2)]\right)+\varphi_2(u_3)\right) \quad (3)$$

where u_1, u_2, and u_3 represent the marginal distributions of Hs, Tp, and S, respectively and φ_1 and φ_2 are the generator functions of the Frank and Gumbel Copulas. It is noted that other and more flexible (yet also more complex) methods to construct higher-dimensional Copula models are available (see e.g. Berg & Aas, 2007).

The results from the trivariate statistical analysis are shown in Fig. 6 as red isolines for a storm surge water level (S) exceeding a value of 205 cmNN (the pdf's of all three marginal parameters are

shown in separate subplots). This decision was made in order to be able to present the results (in the trivariate case the isolines become an isosurface) and compare them with the results from the bivariate analysis of the pair Tp | Hs alone. The results show that the isolines move to the lower left for the trivariate case, as it could be expected when a third variable is added to the model. However, for smaller joint exceedance probabilities the differences between the results from the bivariate and trivariate analyses are negligible. This is because we have chosen that S exceeds a value of 205 cmNN; such an event has a (univariate) return period of ~7.5 years. For the smaller joint exceedance probabilities, where Hs and Tp become high, it is very likely (due to the dependence structure between the different variables) that S is also high and therefore the difference between the isolines from the trivariate and bivariate analyses are small.

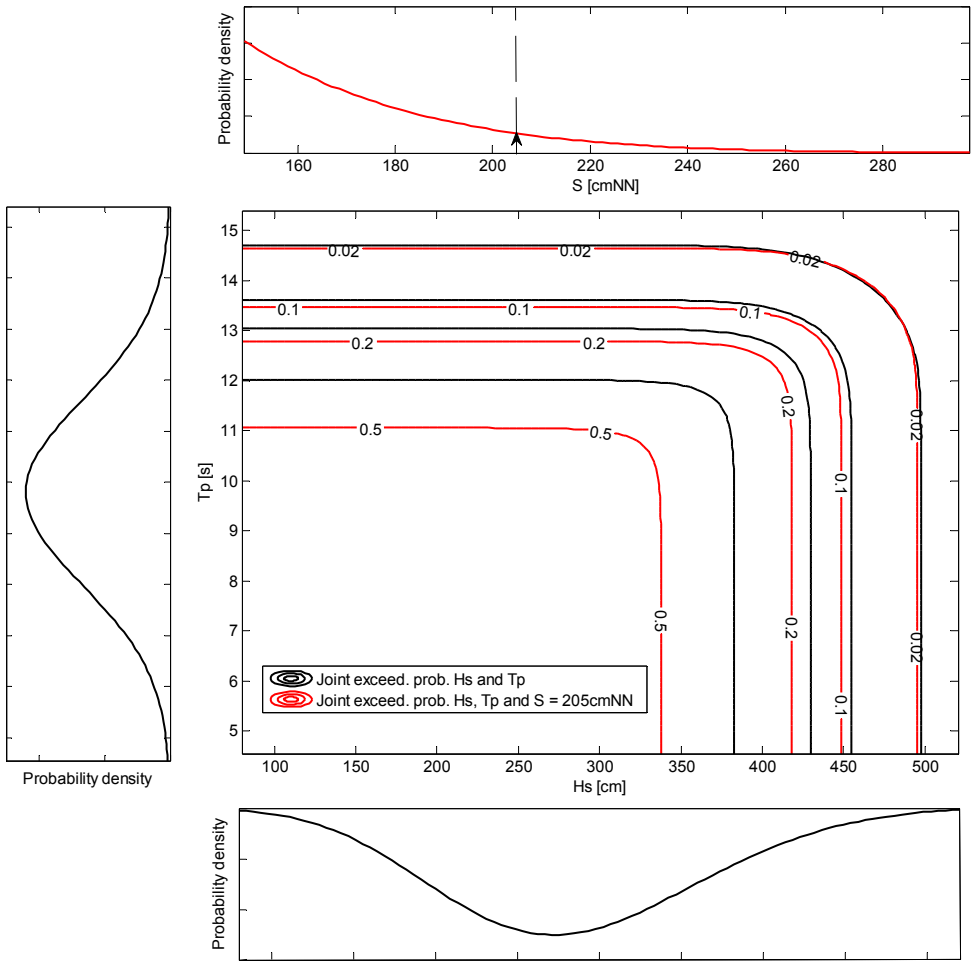


Figure 6: Comparison of the results from a bivariate statistical analysis of the variables Hs and Tp and the results from a trivariate statistical analysis (with a fully nested Archimedean Copula model) of the variables Tp, Hs and S (with S = 205 cmNN for presenting purposes). The probability density functions of the marginal parameters are shown in separate subplots

5 SUMMARY AND CONCLUSIONS

This paper had the intention to make coastal engineers aware of the fact that flexible multivariate models, such as Copulas, are available and have been proven to be applicable to solve a vast number of multivariate problems. We provided a very short and basic introduction to the Copula concept and the interested reader is refereed to text books, such as by Nelson (1999), Joe (1997), or Salvadori *et al.* (2007) providing a much more thorough introduction into the topic.

In Sect. 3 several examples were discussed where Copulas can be useful for common coastal engineering applications, including the analysis of hydraulic loading factors for coastal/offshore infrastructure and sandy beaches, or the calculation of failure probabilities of existent structures

such as dikes, dunes, or sea walls. Finally, in the previous Sect. 4 we provided an example of analyzing a multivariate data set consisting of simultaneously measured storm surge water levels, wave heights, and wave periods using various bivariate and a trivariate fully nested Archimedean Copula. Joint exceedance probabilities (here, the probability that all variables exceed a certain threshold at the same time) were calculated for the variable combinations displayed in Fig. 7.

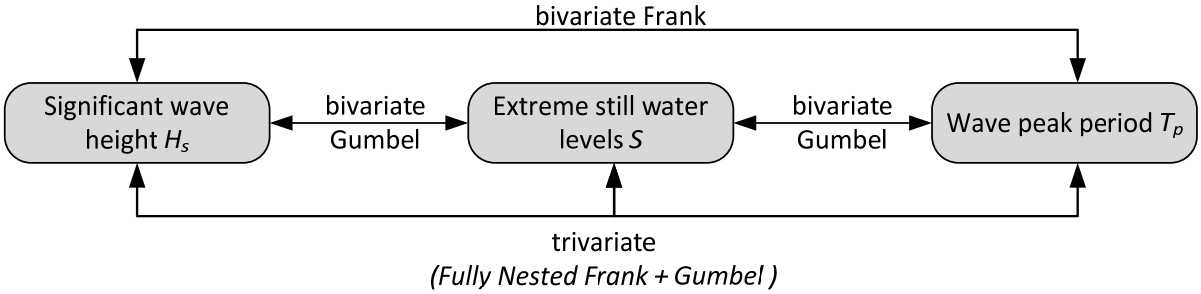


Figure 7: Bivariate and trivariate statistical analyses performed in this study

The discussion provided above and the example given here, clearly demonstrate that multivariate statistical models, and especially Copulas, can be exploited by coastal engineers and applied to various common problems that arise for example while designing coastal and offshore infrastructure, performing coastal flood risk analyses, investigating beach erosion, or calculating failure probabilities.

6 REFERENCES

Berg, D. and Aas, K. (2007): *Models for construction of multivariate dependence: A comparison study*, Europ. J. Finance http://www.informaworld.com/smpp/title~db=all~content=t713684998~tab=issueslist~branches=15 - v1515 (7–8), pp.639–659.

Berg, G. (2009): *Copula goodness-of-fit testing: an overview and power comparison*. Europ. J. Finance, 15 (7–8), pp.675–701.

Coles, S. (2001): *An Introduction to Statistical Modeling of Extreme Values*, Springer, ISBN 1-85233-459-2.

Corbella, S. and Stretch, D.D. (2012): *Predicting coastal erosion trends using non-stationary statistics and process-based models*. Coastal Engineering 70, pp.40–49.

Dixon, M.J. and Tawn, J.A. (1995): *Extreme sea-levels at the UK A-class sites: optimal site-by-site analysis and spatial analyses for the East Coast*, POL Internal Document Number 72. (available from http://www.pol.ac.uk/ntslf/pdf/id72.pdf)

Favre, A.-C., El Adlouni, S., Perreault, L., Thiémonge, N. and Bobée, B. (2004): *Multivariate hydrological frequency analysis using copulas*, Water Resour. Res. 40, W01101. doi:10.1029/2003WR002456.

Genest, C. and Favre, A.-C. (2007): *Everything you always wanted to know about copula modeling but were afraid to ask*, J. Hydrol. Eng. 12 (4), pp.347–368.

Genest, C., Rémillard, B. and Beaudoin, D. (2009): *Goodness-of-fit tests for copulas: A review and a power study*, Insurance Math. Econom 44, pp.199–213.

Greenwood, P.E. and Nikulin, M.S. (1996): *A Guide to Chi-Squared Testing*, John Wiley and Sons Inc., New York, ISBN: 047155779X, 280p.

Gringorten, I. I. (1963): *A plotting rule for extreme probability paper*. J. Geophys. Res. 68 (3), pp.813–814.

Hutchinson, T.P. and Lai, C.D. (1991): *The engineering statistician's guide to continuous bivariate distributions*, Rumsby Scientific Pub., ISBN: 0646024132, 346p.

Joe, H. (1997): *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.

Klein, B., Pahlow, M., Hundecha, Y., Gattke, C. and Schumann, A. (2008): *Probabilistic Analysis of Hydrological Loads to Optimize the Design of Flood Control Systems*, 4th International Symposium on Flood Defence: Managing Flood Risk, Reliability and Vulnerability, Toronto, Canada.

Kortenhaus, A., Oumraci, H., Weissman, R. and Richwein, W. (2002): *Failure mode and fault tree analysis for sea and estuary dikes*. Proc. of the Int. Conf. on Coastal Engineering, Cardiff – Wales.

Nelsen, R.B. (1999): *An introduction to copulas*. Lecture Notes in Statistics, 139, Springer, New York.

Salecker, D., Gruhn, A., Schlamkow, C. and Fröhle, P. (2011): *Parameterization of storm surges as a basis for assessment of risks of failure for coastal protection measures*, Proc. of the 5th International short conference on applied coastal research (SCACR), Aachen, Germany.

Salvadori, G., De Michele, C., Kottegoda, N.T. and Rosso, R. (2007): *Extremes in nature: An approach using copulas*, ISBN: 1402044143, Springer.

Salvadori, G., De Michele, C. and Durante, F. (2011): *On the return period and design in a multivariate framework*, Hydrol. Earth Syst. Sci. 15, pp.3293-3305, doi:10.5194/hess-15-3293-2011.

Salvadori, G., Tomasicchio, G.R. and D'Alessandro, F. (2013): *Multivariate approach to design coastal and off-shore structures*. In: Conley, D.C., Masselink, G., Russell, P.E. and O'Hare, T.J. (eds.), Proceedings 12th International Coastal Symposium (Plymouth, England), Journal of Coastal Research, Special Issue No. 65, pp.386-391, ISSN 0749-0208.

Serinaldi, F. and Grimaldi, S. (2007): *Fully nested 3-copula: procedure and application on hydrological data*, J. Hydrol. Eng. 12 (4), pp.420–430.

Tawn, J.A. and Vassie, J.M. (1989): *Extreme sea levels: The joint probabilities method revisited and revised*, Proc.-Inst. Civ. Eng. 87, pp.429– 442.

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

Wahl, T. and Jensen, J. (2011): *Statistical methods to assess the hydrodynamic boundary conditions for flood risk analyses in coastal areas*, Proc. of the Coasts and Ports Conference, Perth, Australia.

Wahl, T., Mudersbach, C. and Jensen, J. (2012a): *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, doi:10.5194/nhess-12-495-2012.

Wahl, T., Mudersbach, C. and Jensen, J. (2012b): *Statistical assessment of storm surge scenarios within integrated risk analyses – Results of the XtremRisK project*, Proceedings of the 2nd European Conference on FLOODrisk Management, Rotterdam, Netherlands.

Zhong, H. van Overloop, P.-J. and van Gelder, P. (2013): *A joint probability approach using a 1-D hydrodynamic model for estimating high water level frequencies in the Dutch Lower Rhine Delta*, Nat. Hazards Earth Syst. Sci. 13, pp.1841–1852.

Linear and nonlinear modelling for nonstationary annual maximum frequency analysis of storm surges

Galiatsatou P.[1], Prinos P., Anagnostopoulou Chr. and Vasiliadis L.

[1]Division of Hydraulics and Environmental Research, Department of Civil Engineering, A.U.Th, Thessaloniki, Greece, Email: pgaliats@civil.auth.gr

Abstract

In this study the GEV distribution function is used to assess nonstationarity in annual maximum storm surge events simulated from a high resolution storm surge model formulated for the Greek seas at large and forced with simulated data of wind and pressure fields from a RCM. The GEV distribution parameters are specified as functions of time-varying covariates and estimated using a GEV-CDN model proposed by Cannon (2010). Model parameters are estimated via the GML approach using the quasi-Newton BFGS optimization algorithm, and the appropriate GEV-CDN model architecture for each location is selected by fitting increasingly complicated models and choosing the one that minimizes appropriate cost-complexity model selection criteria. For each location examined, different formulations are tested with combinational cases of stationary and nonstationary parameters of the GEV distribution, linear and nonlinear architecture of the CDN and combinations of the input covariates. The wind field, namely wind speed and wind direction and the magnitude of low pressure systems in the form of sea level pressure anomalies, are used in the present work as covariates for the different nonstationary GEV-CDN models. Time is also utilized as one of the covariates in the nonstationary models.

1 INTRODUCTION

One of the main decisions to be made in coastal hydraulics is to estimate design floods for safety purposes. Therefore, the significance of adequately modelling the most likely estimates of extreme marine events, namely high wave conditions and high sea levels (astronomical tide and storm surge) offshore and transformed to nearshore, associated with high return periods together with the uncertainty that complements these estimates is fully perceived. It is also well known that most marine signals exhibit phenomena of nonstationarity attributed mainly to natural climate variability. Global climate change, associated with changes in both the frequency and magnitude of extreme events, is also expected to contribute significantly to the nonstationary behaviour of extreme marine climate.

Extreme Value Theory (EVT) is commonly used to analyse extreme marine events and to extrapolate to levels more extreme than those observed. The inception that hydraulic processes are nonstationary on time scales that are relevant to extreme value analysis, strongly supports the use of extreme value techniques that can account for such phenomena, undermining the utilisation of typical stationary methods. Nonstationarity is therefore incorporated in extreme value models in the form of covariates leading to relatively unbiased quantile estimates, compared to the results of stationary models fitted to heterogeneous samples ignoring covariate effects.

Covariate modelling of the parameters of extreme value models has started to receive considerable attention during the last decades. Coles (2001) presents some simple nonstationary extreme value models with time-varying parameters, which are fitted to extreme storm surge data from the English coast on the North Sea. Anderson *et al.* (2001) examine the spatial and temporal variability of significant wave height measurements and its effect on the estimates of extreme events and discuss the detection of trends in such extremes. Gaetan & Grigoletto (2004) present a semiparametric approach for smoothing sample extremes with dynamic models and use it to assess trends in temperature extremes. Butler (2005) fits parametric polynomial models to estimate long-term trends in storm surge data from the North Sea, while Butler *et al.* (2007) investigate and simulate long-term

trends in storm surge data in the southern and central part of the North Sea, utilizing nonparametric techniques. Coles & Tawn (2005) simulate long-term trends in storm surge data from the North Sea within the Bayesian framework. Stefanakos & Athanassoulis (2006) present a method of calculating return levels from nonstationary time series, where the significant wave height series is expressed as a function of deterministic time dependent periodic functions with a period of one year and of a zero-mean stationary stochastic process. Huerta & Sanso (2007) analyse extreme values of daily ozone levels in the City of Mexico using an extreme value distribution with its parameters defining a space-time structure. The temporal component of the structure is defined through a Dynamic Linear Model (DLM) or state space representation. Yee & Stephenson (2007) introduce the classes of vector generalized linear and additive models which allow all parameters of extreme value distributions to be modelled as linear or smooth functions of covariates. Méndez *et al.* (2008) introduce a time-dependent version of the POT approach for extremes of significant wave height, conditioning to the duration of the storm and accounting for seasonality and apply it to specific reanalysis time series and NOAA buoy records, while Menéndez *et al.* (2009) develop a time-dependent GEV model for monthly significant wave heights maxima, using harmonic functions to represent its parameters. Villarini *et al.* (2009) develop a framework for flood frequency analysis of annual peak records in an urban drainage basin, based on the Generalized Additive Models for Location, Scale and Shape parameters (GAMLSS), while Villarini *et al.* (2010) use the GAMLSS to model extremes from a long record of seasonal rainfall and temperature in Rome highlighting the role of covariate analysis with selected teleconnection indexes in modelling and prediction of extremes. Cannon (2010) proposes a GEV-CDN model which can be used to perform nonlinear nonstationary analysis of hydrological extremes. The parameters of the GEV distribution are specified as a function of covariates using a Conditional Density Network (CDN), which is a probabilistic extension of the multilayer perceptron neural network. Galiatsatou & Prinos (2011) combine the continuous wavelet transform and a point process extreme value approach to perform a more reliable simulation of extreme wave height events in selected locations of the Aegean Sea.

In this study the Generalized Extreme Value (GEV) distribution is used to assess nonstationarity in annual maximum storm surges for selected locations in the N. Aegean Sea. Storm surge extremes result from a high resolution storm surge model formulated for the Greek seas at large. The GEV distribution parameters are specified as functions of time-varying covariates and are estimated using the Conditional Density Network (CDN), as proposed by Cannon (2010). Model parameters are estimated via the Generalized Maximum Likelihood (GML) approach. The appropriate GEV-CDN model architecture for each one of the studied locations is selected by fitting increasingly complicated models with combinational cases of stationary and nonstationary parameters of the GEV distribution, linear and nonlinear architecture of the CDN and combinations of the input covariates and finally by choosing the one that minimizes appropriate cost-complexity model selection criteria. The initiating mechanisms of storm surges, namely the wind (wind speed and wind direction) and atmospheric pressure fields resulting from a Regional Climate Model (RCM), as well as the time, are utilized as covariates in the nonstationary GEV-CDN models.

2 THE GEV-CDN MODEL

Extreme Value Theory is considered to be a robust framework for analysing the tail behaviour of distribution functions. Univariate EVT includes models for block maxima, as well as models for exceedances over appropriately defined thresholds, known as peaks over threshold (POT) models. The former are included within the Generalized Extreme Value (GEV) family of distributions (Jenkinson, 1955). The cumulative distribution function of the GEV with parameters μ (location), σ (scale) and ξ (shape) is given by:

$$G(x) = \exp[-\{1+\xi\frac{(x-\mu)}{\sigma}\}^{-1/\xi}],\ \xi \neq 0 \quad \text{and} \quad 1+\xi\frac{(x-\mu)}{\sigma} > 0 \tag{1}$$

Nonstationary processes have characteristics that change systematically through time. Within the EV modeling framework, the parameters of the GEV can be expressed as a function of covariates. Therefore, the extremal behaviour of the studied variable can be related to that of another or to a number of other variables, referred to as covariates. Time itself or any other time-varying variable having some form of impact on the studied phenomenon can be examined as a possible covariate within a nonstationary context. Estimates of extreme quantiles with exceedance probability *p*,

considering nonstationarity can be derived as follows:

$$z_{p_t} = \mu_t - \frac{\sigma_t}{\xi_t}[1-\{-\log(1-p)\}^{-\xi_t}] \quad (2)$$

where μ_t, σ_t >0 and ξ_t are time dependent GEV parameters.

The method of maximum likelihood (Coles, 2001) is commonly used for parameter estimation because of its simplicity and its efficiency when the sample size is sufficiently large. However, for small sample sizes the ML estimates can be quite unstable, especially compared to other estimation techniques. To solve the problem of divergence of the ML estimator, Martins and Stedinger (2000) introduced the use of a prior distribution for the shape parameter of the GEV model, such that ξ is forced to take its most realistic values eliminating potentially invalid values of the parameter. This method, known as the Generalized Maximum Likelihood (GML) estimation, utilizes a Beta distribution as prior for the shape parameter. The penalized likelihood function of the GML is given by:

$$L_n^{(p)} = \prod_{t=1}^{n} f(x_t|\mu_t,\sigma_t,\xi_t)\pi(\xi_t) \quad (3)$$

In the present study, the prior distribution of the shape parameter, ξ, of the nonstationary GEV distribution is $\pi(\xi)$~ Beta (p = 2, q = 3.3). This penalty results in a broader probability density function, with a mode of approximately -0.2 and approximately 90% of its probability mass concentrated between -0.4 and 0.2 (Cannon, 2010). This prior was selected taking into account that the extreme distribution functions for most marine variables have short tails, characterised by a negative value of the estimator for the shape parameter.

The main disadvantage of using parametric models, such as simple linear or log-linear functions, for the GEV parameters is that the dependence of the parameters on covariates should be defined *a priori* and that such models are often neither appropriate nor realistic. On the other hand, the use of semiparametric and nonparametric approaches needs *a priori* specification of the form of interactions between covariates, which also depends on the subjective judgement of the modeller. The proposed GEV-CDN model, developed by Cannon (2010), can be used to perform nonstationary GEV analysis, overcoming the above mentioned pitfalls of both parametric and nonparametric models. In the GEV-CDN model, the parameters of the GEV distribution are specified as a function of covariates using a conditional density network (CDN) (Bishop, 2006; Cawley *et al.*, 2007), which is a probabilistic extension of the standard multilayer perceptron (MLP) neural network. Due to the flexibility of the neural network architecture, the model is capable of representing a wide range of nonstationary dependencies of the studied variables on a number of covariates. The input layer nodes of the neural network are the covariates x_{it} for i=1,..I, while there are three output layer nodes corresponding to the parameters μ_t, σ_t and ξ_t of the nonstationary GEV distribution function. Input and output layer nodes are connected via J-hidden layer nodes. Input and hidden layer nodes are linked using weights $w_{ji}^{(1)}$, while weights $w_{kj}^{(2)}$ are used to connect hidden and output layer nodes. Weight values associated with individual nodes, known as biases, are added to the hidden layer nodes, $b_j^{(1)}$, as well as to the output layer nodes, $b_k^{(2)}$. The output from the j^{th} hidden-layer node is given by:

$$h_{j_t} = m\left(\sum_{i=1}^{I} x_{i_t} w_{ji}^{(1)} + b_j^{(1)}\right) \quad (4)$$

where $m(\ )$ is the hidden layer activation function. For linear nonstationary GEV-CDN models, the identity function is used for $m(\ )$, while for nonlinear nonstationary models the activation function is the hyperbolic tangent function, tanh( ). The value of the k^{th} output from the model is given by:

$$o_{kt} = \sum_{i=1}^{J} h_{jt} w_{kj}^{(2)} + b_k^{(2)} \quad (5)$$

The parameters of the non-stationary GEV distribution function can be obtained from equation (5) using different activation functions, $g_k(\ )$. More details on the architecture of a GEV-CDN model can be found in Cannon (2010).

In the present study three different categories of candidate GEV-CDN models are implemented for the studied variables. The first one includes only the stationary GEV-CDN model, considering all the parameters of the distribution to be constant with time. The total number of adjustable parameters for this model is P=3. The second category includes the GEV-CDN models, for which the mapping is linear. For these models, the activation function $m(\ )$ of equation (4) is considered to be the identity function. For simplicity, the number of hidden layer nodes is set equal to one. The third category of models includes the nonlinear GEV-CDN models. The hidden layer activation function for these models is the hyperbolic tangent function, tanh( ). For a number of I covariates and J hidden layer nodes, the total number of adjustable parameters is $P= J(I+3)+3$. In the present work, the number of hidden layer nodes ranges between one and three.

All candidate models are trained by a Broyden - Fletcher - Goldfarb - Shanno (BFGS) quasi-Newton optimisation algorithm to minimise the GML cost function. To avoid convergence of the algorithm to local minima, the optimization algorithm is run using fifty random restarts. Parameters (weights and biases) of the neural network associated with the maximum GML over the random restarts are selected as the final parameters. The appropriate GEV-CDN architecture among all candidate models is selected by means of the minimization of the Akaike Information Criterion with small sample size correction (AICc) (Hurvich & Tsai, 1989). The Bayesian information criterion (BIC) (Schwarz, 1978) is also utilized as a complimentary cost-complexity model selection criterion.

To ensure a good generalisation of the methodology and to obtain more reliable parameter estimates for the selected models, more computationally intensive methods are also used in the present work. Therefore, the technique of bagging with ensembles is applied. The method trains multiple models on different samples (data splits) and averages their predictions, improving the accuracy of a selected model. In each iteration step, a random sampling with replacement from the training dataset is performed to create a number of samples and then the selected models are trained on these samples. For each studied case, two hundred ensembles are created from the original training datasets using the method of bootstrap resampling with replacement. The multiple datasets are trained for the selected GEV-CDN models to contribute to building multiple classifiers, which are finally combined.

To estimate confidence intervals for the parameters and the quantiles of the nonstationary GEV distribution function, bootstrap-based techniques are utilized. More specifically, the parametric bootstrap approach is used. The procedures followed within a parametric bootstrap approach include: a) The fitting of a nonstationary GEV model to the data, b) The random sampling from the fitted distribution function, c) The fitting of a nonstationary GEV model to the bootstrapped samples and d) The estimation of parameters and quantiles from the bootstrapped samples. Five hundred bootstrap samples are generated for the selected GEV-CDN models. For each test case, fifty trials are run to fit the selected architecture to the bootstrapped samples. Quantiles and their respective confidence intervals are estimated by substituting the estimated parameters from the bootstrapped samples in equation (2) and calculating percentiles of the respective distributions.

3 DATASETS

The methods and techniques of the present work are implemented to storm surge extremes available at two selected areas of the N. Aegean Sea, namely locations close to the port of Alexandroupolis and Chios. The storm surge values (m) for each selected location are annual maxima and can therefore be considered independent and identically distributed (iid) random variables to be simulated by means of a GEV distribution function. Storm surge extremes result from simulations from a high resolution two-dimensional model of hydrodynamic ocean circulation formulated for the Greek seas at large (Krestenitis *et al.*, 2011). The model validation and calibration has been performed using historic sea level height data from several in-situ stations, located along the Greek coasts (hindcast simulations). The dataset used in the present work covers a period of 149 years, namely from 1952 to 2100.

The forcing of the model consists of simulated data of wind and pressure fields derived from a Regional Climate Model (RCM), RegCM3. RegCM3, was built upon the NCAR-Pennsylvania State University (PSU) Mesoscale Model version 4 (MM4) (Dickinson *et al.*, 1989). The spatial resolution of the model is 25x25 km. Its future projections are forced by the A1B SRES scenario (Jacob *et al.*, 2007). The wind field, namely wind speed and wind direction and the magnitude of low pressure systems in the form of sea level pressure anomalies, are used in the present work as covariates for the different nonstationary GEV-CDN models. Time is also used as one of the covariates in the models. It should be noted that utilized values of wind speed, wind direction and sea level pressure anomalies occur on the same day with the maximum annual surge.

For each location under study, different combinations of covariates are tested and therefore the nonstationary models created range from the most simple ones containing only one covariate to the most complex ones with four covariates, namely sea level pressure, wind speed, wind direction and time. For the models that depend on both the wind speed and the wind direction, the two components of wind speed (east-west component - Windu and north-south component - Windv) are utilized in the model adjustment procedures.

4 RESULTS AND DISCUSSION

For both locations considered a maximum of four covariates is included in the models constructed. Twenty-nine candidate GEV-CDN models are tested, to select the most appropriate GEV-CDN architecture for the studied datasets and the different combinations of the selected covariates. The examined models include a stationary one, seven linear models with non-stationary location, scale or shape or possible combinations and twenty-one nonlinear models with non-stationary location, scale or shape or possible combinations using J = 1, 2, 3 hidden layer nodes.

For storm surge annual maxima for both locations considered, the highest correlation coefficients in absolute values are observed for the sea level pressure anomalies (ρ = -0.38 for Alexandroupolis and ρ = -0.46 for Chios) and the time covariate (ρ = -0.23 for Alexandroupolis and ρ = -0.36 for Chios). Positive correlation coefficients are estimated for the covariate of wind speed, namely ρ = 0.15 for Alexandroupolis and ρ = 0.09 for Chios. Considering wind direction, appreciable correlation coefficients are estimated for the east-west component of wind speed for the location of Alexandroupolis (ρ = 0.28) and for both the east-west and the north-south components for the location of Chios (ρ = 0.13 and ρ = 0.18, respectively).

From all candidate models and possible covariates, Table 1 presents for each location some of the cases examined (the stationary model and the top three covariate/model combinations in terms of the penalty functions analysed in Section 2), together with their resulting AICc and BIC values.

Table 1: Model performance of GEV-CDN models for Alexandroupolis and Chios

Location	Model	Covariates	Hidden Layers	AICc	BIC
Alex/polis	Stationary	-	-	-88.24	-79.39
	Nonlinear with nonstationary μ and σ	Time + Slp + Windu + Windv	2	-165.50	-119.11
	Nonlinear with nonstationary μ and σ	Time + Slp + Windu + Windv	3	-162.06	-99.64
	Nonlinear with nonstationary μ and σ	Slp + Wind Speed	3	-145.78	-96.98
Chios	Stationary	-	-	-70.60	-61.75
	Nonlinear with nonstationary μ and σ	Time + Slp + Windu + Windv	3	-154.19	-92.77
	Nonlinear with nonstationary μ and σ	Slp + Windu + Windv	3	-152.24	-96.43
	Nonlinear with nonstationary μ and σ	Time + Slp + Windu + Windv	2	-142.55	-96.15

Based on the AICc and BIC criteria, the best GEV-CDN model for storm surge extremes at the location of Alexandroupolis is the nonlinear model with two hidden layer nodes with location and scale parameters which are nonstationary functions of the time, sea level pressure anomalies, wind speed and wind direction variables (AICc = -165.50, BIC = -119.11). The values of the AICc and the BIC for the selected model are significantly lower than the ones of the other tested covariate/model combinations. For storm surge extremes in Chios the nonlinear model with three hidden layer nodes with location and scale parameters which are nonstationary functions of sea level pressure anomalies, east-west and north-south components of wind speed is recommended as the best model (AICc = -152.24, BIC = -96.43). The model presents the second lowest AICc and BIC values among the tested ones, but considering the minimisation of both criteria it is judged to perform better than the others.

For the location of Alexandroupolis and for the selected nonlinear GEV-CDN model (with two hidden layers and nonstationary location and scale parameters varying with time, sea level pressure, east-west and north-south wind components), contour plots of relationships among some of the covariates and the τ = 0.99 storm surge quantiles are presented in Figure 1. The storm surge quantiles are estimated using bagging techniques. It should be noted that black dots indicate values of the utilized covariates.

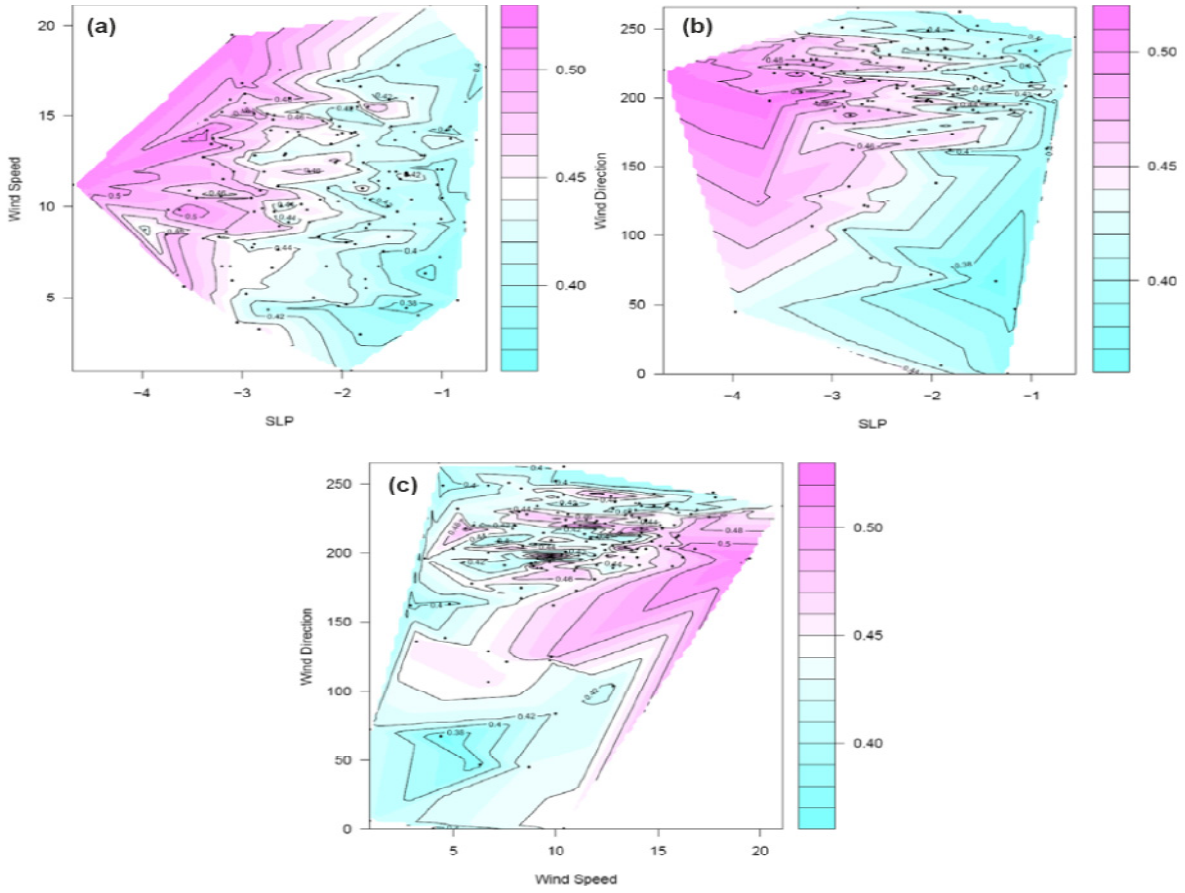


Figure 1: Contour plots of relationships among storm surge τ = 0.99 quantile and (a) wind speed and sea level pressure anomalies, (b) wind direction and sea level pressure anomalies and (c) wind direction and wind speed for Alexandroupolis for the selected GEV-CDN model

From Figure 1, it is evident that the adjusted model contains more than two covariates and therefore the relationships among them are quite difficult to interpret by means of two dimensional plots. However, it can be noted that large surge values occur for sea level pressure anomalies ranging from 3 ~ 4.5 standard deviations below the mean, for wind speeds of more than 9 m/s and for wind directions ranging between 150°-230°. It can also be noted that for negative phases of sea level pressure anomalies, the covariate of pressure seems to have a very significant impact on the estimated storm surge quantiles. It can also be observed that for sea level pressure anomalies ranging from 2 ~ 3 standard deviations below the mean, the relationships between the covariates of

wind speed and slp become more perplexed. In the wind speed-wind direction space, the largest storm surges are estimated for directions ranging between 150 - 220°. For directions 100 - 150° and for wind speeds in the range of 3 - 7 m/s medium to large 0.99 surge quantile values are generated. The contour plots of the relationships between time and the other three covariates (not shown here for the sake of brevity) reveal the highest surge values for the time period 1952-2050. For the last fifty years of the time period considered the storm surge quantiles seem to have a nonlinear decreasing trend. But it should be emphasized that caution is required in the interpretation of the GEV-CDN model results, due to the small number of samples in some regions of the space of covariates.

When a more simple GEV-CDN model with two covariates, namely the sea level pressure and the wind speed, is utilized for storm surge maxima in Alexandroupolis, it is easier to draw inference on interactions between covariates, as well as relationships among the covariates and the GEV distribution parameters. For Alexandroupolis the nonlinear GEV-CDN model with three hidden layers and nonstationary location and scale parameters varying with the sea level pressure anomalies and the wind speed is judged to be the third best model according to the selected penalty functions. Figure 2 presents the location parameter of the GEV model and the τ = 0.99 storm surge quantile using bagging techniques for this more simple model. The variations in the τ = 0.99 storm surge quantiles seem to follow the location parameter of the model in the covariate space. The primary influence on storm surge seems to be the sea level pressure, since largest values of the location parameter as well as the τ = 0.99 quantile occur for sea level pressure anomalies more than 3.5 standard deviations below the mean, while the lowest ones occur for less than 2 standard deviations below the mean. The wind speed contributes to large storm surges, only when the sea level pressure anomalies are low enough. The location parameter of the model as well as the presented surge quantiles decrease with increasing sea level pressure for a wide range of wind speed values. The scale parameter of the GEV model (not shown here for the shake of brevity) seems to be solely influenced by sea level pressure anomalies.

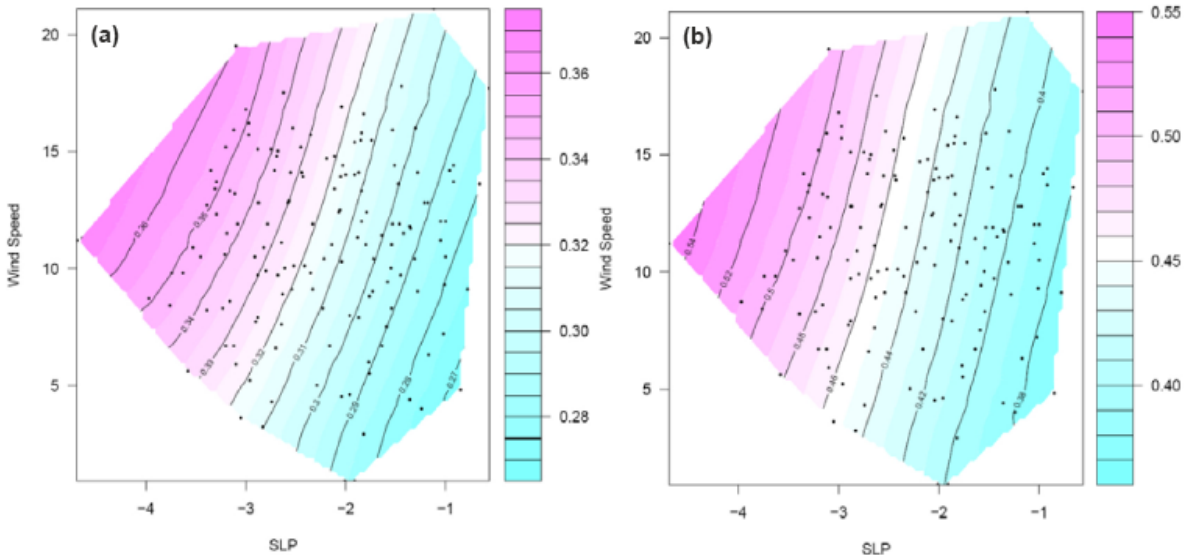


Figure 2: Contour plots of relationships among (a) the GEV location parameter and (b) the τ =0.99 storm surge quantiles and the covariates of sea level pressure anomalies and wind speed

For the location of Chios and for the selected nonlinear GEV-CDN model (with three hidden layers and nonstationary location and scale parameters varying with sea level pressure, east-west and north-south wind components), contour plots of relationships among the covariates and the τ = 0.99 storm surge quantiles are presented in Figure 3. The storm surge quantiles are estimated using bagging techniques. Black dots indicate simulated values of the utilized covariates. The sea level pressure seems again to be the primary influence on storm surge events. The τ = 0.99 surge quantiles decrease with increasing sea level pressure for a wide range of wind speed and wind direction values. Large values of storm surge quantiles occur for sea level pressure anomalies more than 3 standard deviations below the mean even for low enough wind speeds and a for a wide range of wind dimensions. For high wind speeds and wind directions exceeding 250°, the storm surge quantiles seem to be modified showing a decreasing trend with constant sea level pressure

anomalies. Considering wind speed and wind direction, the largest storm surges occur for wind speeds exceeding 10 m/s and wind directions ranging between 140°-280°.

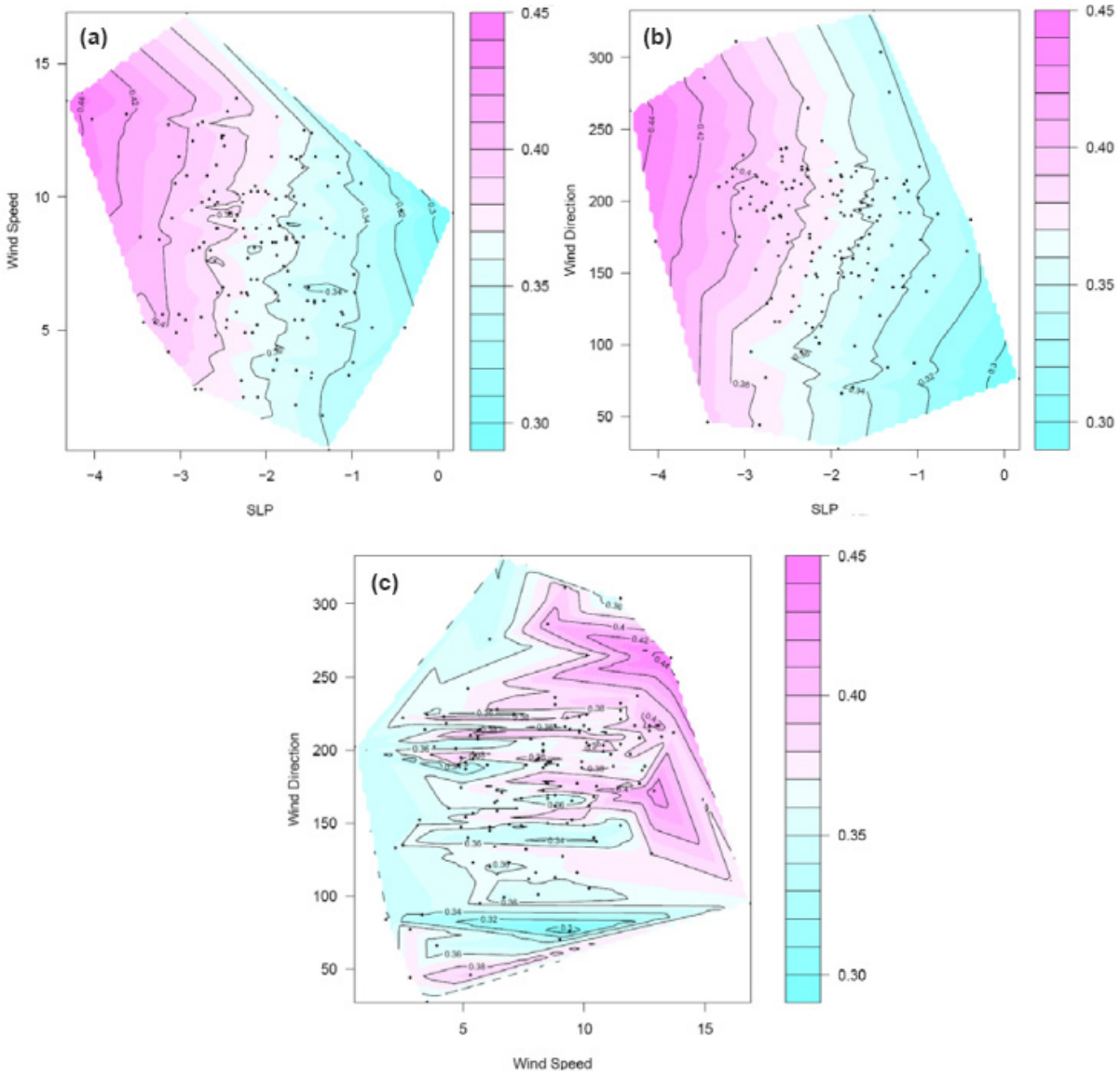


Figure 3: Contour plots of relationships among storm surge τ = 0.99 quantile and (a) wind speed and sea level pressure anomalies, (b) wind direction and sea level pressure anomalies and (c) wind direction and wind speed for Chios for the selected GEV-CDN model

Figure 4 shows τ = 0.99 storm surge quantiles (black solid line) and 95% confidence intervals (blue dashed lines) for the selected nonlinear GEV-CDN models with two hidden layers and nonstationary location and scale parameters varying with time, sea level pressure, east-west and north-south wind components for Alexandroupolis and three hidden layers and nonstationary location and scale parameters varying with sea level pressure, east-west and north-south wind components for Chios. Black dots indicate simulated values of annual maximum storm surge events. In many cases, local maxima for the presented quantiles correspond to years of high annual storm surges, while local minima to years of lower events. High spikes of storm surge quantiles in years with not very high surge events or relatively low values in years with high storm surges are possibly attributed to the combination of covariates and their nonlinear interactions.

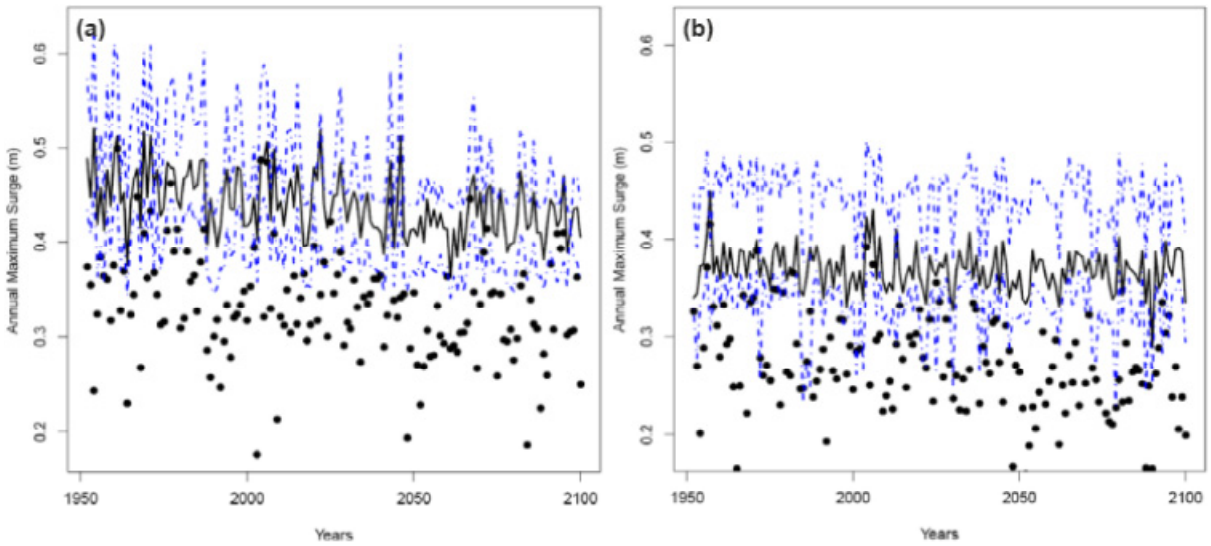


Figure 4: Estimates of τ =0.99 storm surge quantiles with 95% confidence intervals for the selected GEV-CDN models for (a) Alexandroupolis and (b) Chios

5 CONCLUSIONS

In this study the GEV distribution function is used to assess nonstationarity in annual maximum storm surge events simulated from a high resolution storm surge model formulated for the Greek seas at large and forced with simulated data of wind and pressure fields from a RCM. The GEV distribution parameters are specified as functions of time-varying covariates and estimated using a GEV-CDN model. Model parameters are estimated via the GML approach using the quasi-Newton BFGS optimization algorithm, and the appropriate GEV-CDN model architecture for each location is selected by fitting increasingly complicated models and choosing the one that minimizes appropriate cost-complexity model selection criteria. For each location examined, different formulations are tested with combinational cases of stationary and nonstationary parameters of the GEV distribution, linear and nonlinear architecture of the CDN and combinations of the input covariates.

The implemented GEV-CDN models are used to apply nonparametric approaches to nonstationary extreme value analysis of storm surge data to overcome the usual linear assumption of parameter dependence on covariates and have been proven to account for interactions between covariates, as well as between covariates and GEV distribution parameters or quantiles, without *a priori* specification of the form of these interactions. For both locations considered nonlinear GEV-CDN models with nonstationary location and scale parameters provided better fit to the storm surge data. The analysis revealed not only the prevailing impact of sea level pressure on the magnitude of storm surge extremes, but also its interactions with the wind field variables, the wind speed and the wind direction. The best fitted models for both locations considered included both the forcing mechanisms of storm surges. The time covariate was also considered necessary for modelling nonstationarity in one of the two selected locations.

6 ACKNOWLEDGMENTS

This work was conducted within the research project “CCSEWAVS: Estimating the effects of Climate Change on SEa level and WAve climate of the Greek seas, coastal Vulnerability and Safety of coastal and marine structures”. The authors want to thank the Department of Meteorology and Climatology, School of Geology of A.U.Th for providing the data of the wind and pressure fields and the research team of Prof. Yannis Krestenitis from the Division of Hydraulics and Environmental Engineering of the Civil Engineering Department of A.U.Th for providing the storm surge data. The R package GEVcdn by Alex J. Cannon was utilized for the applications described in the paper.

7 REFERENCES

Anderson, C.W., Carter, D.J.T. and Cotton, P.D. (2001): *Wave climate variability and impact on offshore design extremes*, Report for Shell International and the Organization of Oil & Gas Producers.

Bishop, C.M. (2006): *Pattern Recognition and Machine Learning*, Springer-Verlag, New York.

Butler, A. (2005): *Statistical modelling of synthetic oceanographic extremes*, Ph.D. Thesis, Lancaster University, London, UK.

Butler, A., Heffernan, J.E., Tawn, J.A. and Flather, R.A. (2007): *Trend estimation in extremes of North Sea surges*, Appl. Statist 56, pp.395-414.

Cannon, A.J. (2010): *A flexible nonlinear modelling framework for nonstationary generalized extreme value analysis in hydroclimatology*, Hydrol Process 24, pp.673-685.

Cawley, G.C., Janacek, G.J., Haylock, M.R. and Dorling, S.R. (2007): *Predictive uncertainty in environmental modelling*, Neural Networks 20, pp.537-549.

Coles, S. (2001): *An Introduction to Statistical Modelling of Extreme Values*, Springer, London.

Coles, S. and Tawn, J. (2005): *Bayesian modeling of extreme surges on the UK east coast*, Phil Trans R Soc A 363, pp.1387–1406.

Dickinson, R., Errico, R., Giorgi, F. and Bates, G. (1989): *A regional climate model for the western United States*, Clim. Chang 15 (3), pp. 383-422.

Gaetan, C. and Grigoletto, M. (2004): *Smoothing sample extremes with dynamic models*, Extremes, 7, pp.221-236.

Galiatsatou, G. and Prinos, P. (2011), *Modeling non-stationary extreme waves using a point process and wavelets*, Stoch Env Res Risk A 25 (2), pp.165-183.

Huerta, G. and Sanso, B. (2007): *Time-varying models for extreme values*, Environ Ecol Stat 14, pp.285-299.

Hurvich, C.M. and Tsai, C. (1989): *Regression and time series model selection in small samples*, Biometrika, 76, pp.297-307.

Jacob, D., Barring, L., Christensen, O.B., Christensen, J.H., de Castro, M., Deque, M., Giorgi, F., Hagemann, S., Hirschi, M., Jones, R., Kjellstrom, E., Lenderink, G., Rockel, B., van Ulden, A.P. and van den Hurk, B.J.J.M. (2007): *An inter-comparison of regional climate models for Europe:model* performance in present-day climate, Clim Chang 81, pp.31-52.

Jenkinson, A.F. (1955): The frequency distribution of the annual maximum (or minimum) values of *meteorological elements*, Q. J. Roy. Meteor. Soc. 81, pp. 158-171.

Krestenitis, Y., Androulidakis, Y., Kontos, Y. and Georgakopoulos, G. (2011): *Coastal inundation in the north-eastern Mediterranean coastal zone due to storm surge events*, J. Coast Conserv, Springer, 15, pp.353-368.

Martins, E.S. and Stedinger, J.R. (2000): *Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data*, Water Resour. Res 36, pp. 737-744.

Méndez, F.J., Menéndez, M., Luceño, A., Medina, R. and Graham, N.E. (2008): Seasonality and duration in extreme value distributions of significant wave height, Ocean Eng 35, pp.131-138.

Menéndez, M., Méndez, F.J., Izaguirre, C., Luceño, A. and Losada, I.J. (2009): *The influence of seasonality on estimating return values of significant wave height*, Coastal Eng 56, pp.211-219.

Schwarz, G. (1978): *Estimating the dimension of a model*, Ann. Stat 6, pp.461-464.

Stefanakos, C.N. and Athanassoulis, G.A. (2006): *Extreme value predictions based on nonstationary time series of wave data*, Environmetrics 17 (1), pp. 25-46.

Villarini, G., Smith, J.A. and Napolitano, F. (2010): *Nonstationary modelling of a long record of rainfall and temperature over Rome*. Adv. Water Res 33, pp. 1256-1267.

Villarini, G., Smith, J.A., Serinaldi, F., Bates, J., Bates, P.D. and Krajewski, W.F. (2009): *Flood frequency analysis for nonstationary annual peak records in an urban drainage basin.* Adv. Water Res 32, pp. 1255-1266.

Yee, T.W. and Stephenson, A.G. (2007): *Vector generalized linear and additive extreme value models*, Extremes, 10, 1-19.

Impacts of morpho-dynamics and SLR on extreme water level statistics and implications for climate change adaption strategies in coastal Denmark

Carlo Sørensen[1] and Thorsten Piontkowitz[1]

[1]Danish Coastal Authority, Lemvig, Denmark, Email: cas@kyst.dk

Abstract

One advantage of extreme water level statistics is that the natural variability in climate and associated storm surge heights may be described in a simple and useful manner for coastal planning and climate change adaption. SLR, morpho-dynamic changes, land movement and other factors, however, may also have an impact on extreme water levels and on the statistics and their application. The Danish statistics based on tide gauge records from 68 stations are presented together with two case studies. One case contains a comprehensive modelling effort looking into the effect of morpho-dynamic changes on extreme water levels and provides updated statistics based on bathymetry-corrected water levels in the area considered, whereas the other case relates to the use of statistics in climate change adaptation by including land movement and SLR into to a "dynamic" DEM, adjusted to better represent future conditions in relation to coastal flooding hazards.

Apart from considerations on the statistical approach, changes in the above parameters may yield a different picture of coastal flooding hazards and the challenges ahead. In the studied areas SLR only ranks third in importance after subsidence and morpho-dynamic changes at an intermediate time scale. Due attention to these matters is therefore advocated in conjunction to the calculation and communication of extreme water level statistics and climate change adaptation in Denmark.

1 INTRODUCTION

There are indeed challenges related to the production and use of extreme water level statistics from tide gauge measurements. Challenges may relate to the instrumentation (type and design, deployment site, data collection and transmission, data quality and processing, benchmarking and datum corrections) and human attention to these matters; to the choice and application of statistical methods (including, amongst others, aspects of precision and representation in historical data) and to the actual use in relation to present and future coastal flooding hazards. This may seem trivial from a technical and scientific point of view, nevertheless considerations about these aspects are important in the overall assessment, use and communication of the statistics. No matter how good and robust the statistical approach is, the results still depend on the validity of data and thorough consideration about their applicability and use is needed.

Although many aspects may be accounted for in the statistics, we still face limitations in our work due to a lack of data and/or knowledge. Here, following a brief presentation and discussion of the Danish extreme water level statistics, results related to morpho-dynamic changes and land movement are presented from two case studies. They draw attention to the representativity of historical extremes and to the use of the present statistics in climate change adaptation, respectively.

2 THE DANISH EXTREME WATER LEVEL STATISTICS

The Danish extreme water level statistics (Sørensen *et al.*, 2013a) cover 68 locations along the Danish coastline and are based solely on measured extremes in tide gauge records. The tide gauges are owned and maintained by different authorities, e.g. the Danish Meteorological Institute, the Danish Coastal Authority (DCA), the Danish Nature Agency and local municipalities and harbours.

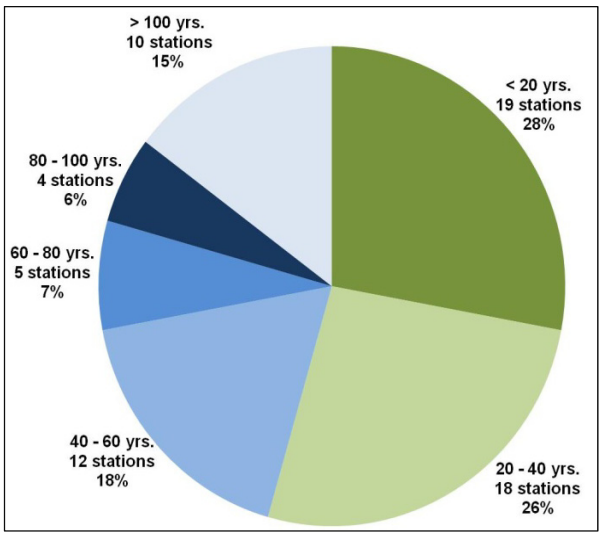


Figure 1: Distribution of data periods (n = 68) used in the extreme water level statistics

The data series in individual statistics range from 11.6 to 139 full years (total 3228 yrs.) where 31 statistics are based on more than 40 years of data, Figure 1.

Some statistics based on only 12 - 15 years of data have been included also to promote into society and to decision-makers the actual usefulness of tide gauge data and to emphasize that its takes at least a decade of sound measurements to have somehow valid time-series for extreme value analysis.

The statistics are widely applied in climate change adaptation, coastal planning and management, in relation to the implementation of the EU Floods Directive and in the assessment of extreme events by the Danish Storm Council where citizens may be eligible for economic compensation for flooding damages on property on occasions exceeding a 20 year event.

2.1 Methodology

Data come in a digital form from a large variety of tide gauges (laser, radar, float, pressure transducer) of which some time series are in a poor condition regarding maintenance, datum, drift etc. Some older data series are kept in an analogue form, whereas from other series only the extreme values exist making a reassessment of data quality difficult or even impossible.

All data are first evaluated, corrected and related to the Danish datum DVR90 (basis in 1990). Extremes, selected above a certain threshold, are then de-trended to mean sea level in the year of occurrence. In practice, a linear interpolation between the former Danish datum of DNN (mean water level around 1890) and DVR90 is used at all stations except in the Limfjord, refer below. Tides are not considered.

In the selection of extremes a time factor of independency between individual events of 36-72 hours is used together with conditions of a normalized water level between events. Further, the independency is checked manually within individual water compartments. Extremes are then compared between neighbouring stations and the spatial distribution of the event is resolved and evaluated.

Statistics are then calculated using a basic POT method for Log-Normal (Wadden Sea and Limfjord) or Weibull distributions according to Sørensen *et al*. (2013a) and Sørensen & Ingvardsen (2007) in which the choice of cut-off level is made manually and thus is liable to some subjectivity. Whereas the extremes fit the distribution very well at some locations, the fit is fair to poor at others due to different meteorological and hydrodynamic conditions being the cause of extreme events.

In order to make the statistics as useful and transparent as possible results from individual stations are presented as return levels for 20, 50 and 100 year events with standard errors, graphics showing the distribution function and a Quantile plot of the fit together with a list of the extremes. Further, the cut-off level, the intensity of occurrence and parameters of the applied distribution are stated together with other station-specific information on data quality, data period etc.

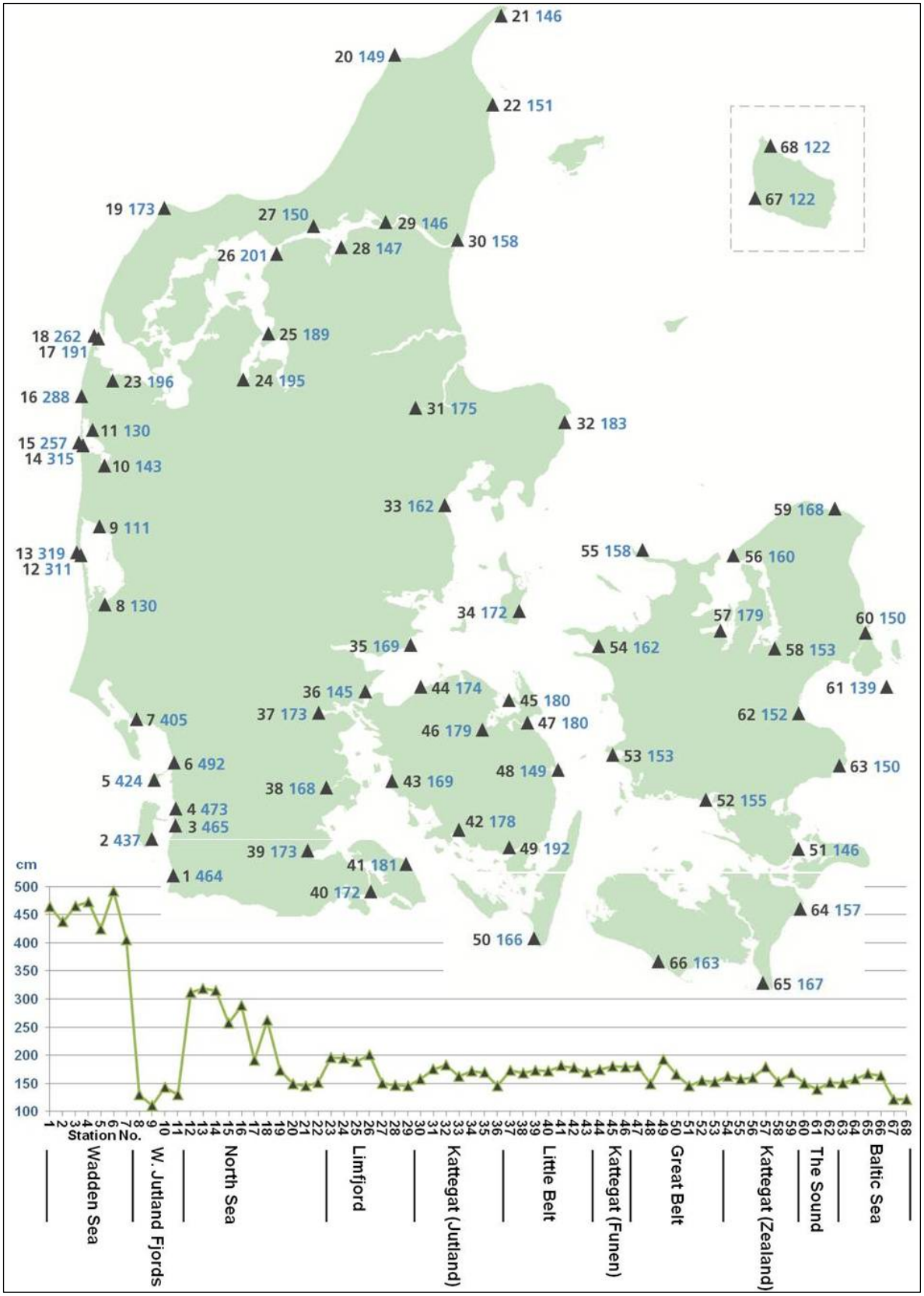


Figure 2: 100 year return heights (blue) in cm (trend-free) for 68 Danish tide gauge stations with the spatial variation in extremes between water compartments sketched (bottom)

2.2 Spatial variations in extremes

Figure 2 shows the 100 year return heights from the tide gauge locations around Denmark. As the results are de-trended to 1990 a couple of centimetres, in general, should be added to the numbers to relate these to DVR90. Any possible acceleration in SLR since 1990 is unaccounted for.

In the Wadden Sea the 100 year return heights are between 405 and 492 cm. Along the North Sea coast extremes are in the order of 300 cm and decreasing to the north to about 150 cm. In most of the remaining water bodies the 100 year return heights are 150 - 200 cm with some variation between stations. In the sluice regulated West Jutland fjords 100 year return heights are below 150 cm as is the case on the island of Bornholm situated in the Baltic Sea (insertion in Figure 2).

In the Wadden Sea and on the North Sea coast there is a fairly straightforward relationship between the meteorological forcing and storm surges occurring at more or less regular intervals, whereas the picture is more complex in the Inner Danish Waters (comprising Kattegat, The Sound and Belts and the Baltic Sea). Very severe events are infrequent and tend to occur only a few times every century in relation to wave phenomena and/or local surges in the narrow parts of the Baltic Sea – North Sea transition. The most extreme water level(s) may thus be considerably higher than the remainder based on “normal” bad weather conditions. The statistics’ representation of a 100 year event may thus be difficult to evaluate and correspondingly may yield a poor fit in the statistics. For example did the surge levels in the 1 - 2 November 2006 event by far exceed any other registration even in some long (> 100 yrs.) tide gauge series. Adding to this, as the water level excursions are smaller in the Inner Danish Waters, the differences in both relative and absolute values between, say, a 20 year and a 100 year event are also small. This, of course, calls for caution when interpreting and using the statistics for planning purposes.

The statistics do however show a consistent distribution function pattern within individual water compartments as exemplified here from the Wadden Sea, Figure 3, where the lines are almost parallel and vertical displacements between stations to a large extent can be explained by the positions of the tide gauge stations within the Wadden Sea.

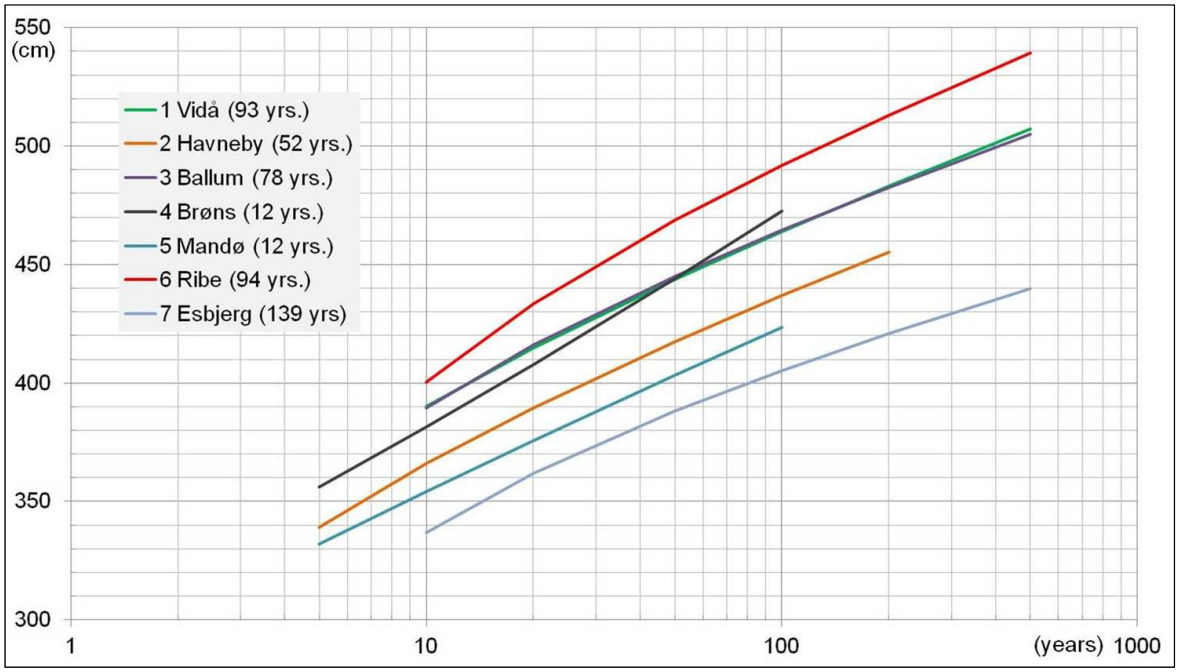


Figure 3: Distribution functions (Log-Normal) of extreme water level statistics at the tide gauge stations (station no., location, and length of time series) in the Wadden Sea. Refer to figure 2 for positions of stations

3 MORPHO-DYNAMICS

Physical and morphological changes, sudden or gradual, may affect the extreme surge levels on a local to regional scale. One consequence of this is that historical extremes may no more be representative in extreme water level statistics.

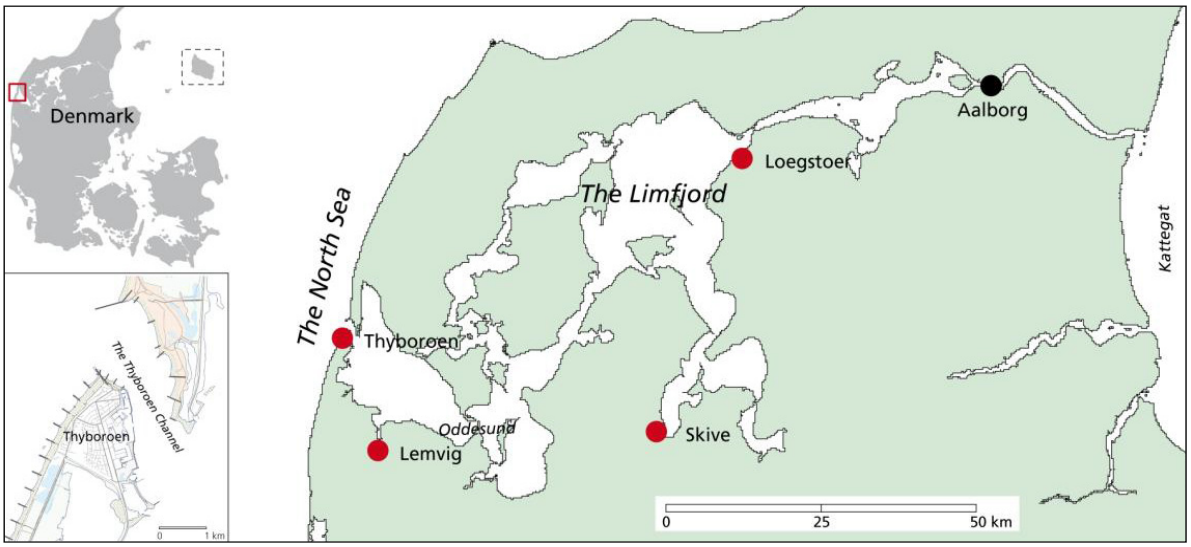


Figure 4: The Limfjord. Insert (bottom left) shows the town of Thyboron and the Thyboron Channel connecting the Limfjord to the North Sea

An apparent gradual increase in extreme events over the last decades in the Limfjord, Figure 4, that could not be explained meteorologically led the DCA to initiate an investigation into the possible causes and magnitudes of change. As the main water exchange in the Limfjord occurs through the Thyboron Channel to the North Sea changes in morpho-dynamic parameters in the channel were in focus from the beginning. Further, the current annual import of sand to the Limfjord amounts to approximately $1*10^6$ m^3 of which the majority settles on the extensive flood tidal delta inside the channel, Figure 5. Please, refer to Christensen (2011a, 2011b), Ingvardsen *et al.* (2012), Knudsen *et al.* (2011, 2012) and Sørensen *et al.* (2013b) for detailed information on methodology, results and recommendations of the investigation.

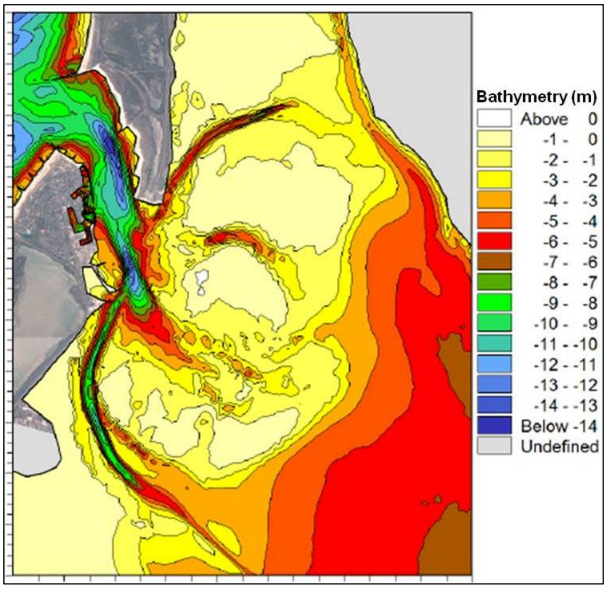


Figure 5: Section of model bathymetry representing 2011 conditions at Thyboron

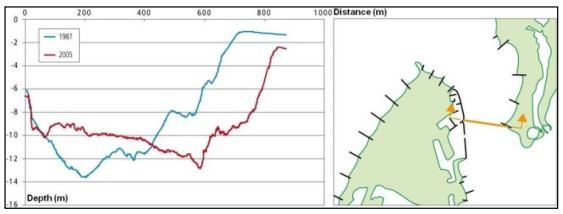


Figure 6: Depths in transect of Thyboron Channel in 1987 (blue) and 2005

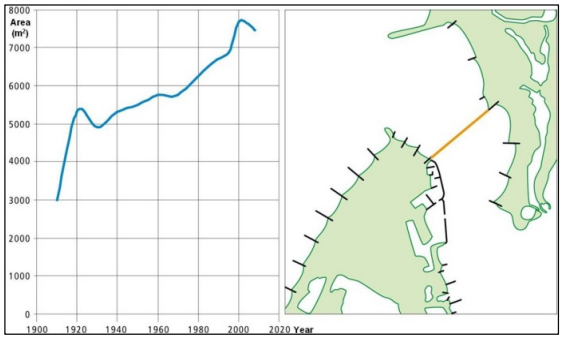


Figure 7: Development in the cross-sectional area of Thyboron Channel since 1910

3.1 Methodology and results

Knudsen *et al.* (2011) found a shift towards east of the deeper parts of the channel since 1987, figure 6, and an increase in the cross-sectional area of the channel from 3000 m^2 to 8000 m^2 since 1910, Figure 7.

A detailed MIKE21 HD model of the entire Limfjord was set up in close collaboration between the Danish Hydraulic Institute (DHI) and DCA and calibrated for 6 recent storm events including data for the entire Limfjord area on bathymetry, wind, water levels, waves and tides. Thorough sensitivity analyses of the importance of waves, wind direction, depths in the channels and on the flood tidal delta, storm intensity, MSL, impulse effects (related to sudden increases in water level in the North Sea) etc were evaluated and showed that changes in the Thyboron Channel, by far, was the factor having the largest impact on storm surge levels in the fjord.

To quantify effects of the changes in the Thyboron Channel on extreme water levels, modelling was carried out using channel bathymetries from 1958, 2005 and 2060, where the 2060 bathymetry was constructed by assuming that the current development of the channel will continue.

The suite of models on calibrated bathymetries was used to establish bathymetry-corrected extreme water level statistics from 1958, 2005 and 2060 for 20 locations in the Limfjord by simulating storms representing 31 high water level events in a 33 year period. The 20 most extreme water levels at each location entered the statistical extreme value analyses to secure a consistent basis for comparison in time and space (the general method is in accordance with the one presented in chapter 2). For the four locations with established extreme water level statistics/tide gauge records at the time of investigation (Thyboron, Lemvig, Skive and Logstor) recalculations based on the bathymetry-corrected water levels were performed on individual storms to minimize the errors on the water level corrections and to better represent the actual effect today (2005) and in the future (2060).

For each of the 20 locations a trend-relation has been established between extreme water levels in the 1958, 2005 and 2060 bathymetries as exemplified by Lemvig in Figure 8. The representation of an extreme water level back and forth in time can then be made by interpolation of the trends in relation to the year of occurrence.

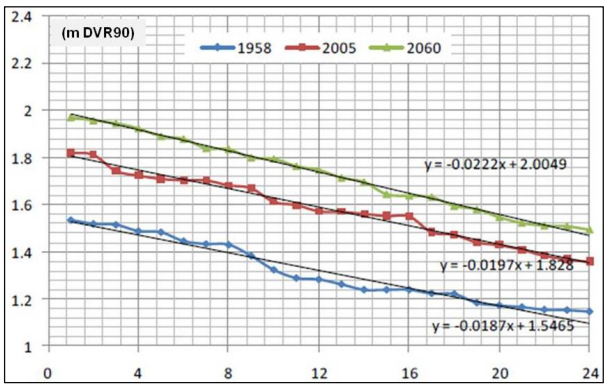


Figure 8: Trend-relations for the modelled storms in Lemvig ranked by surge levels

In general, it was found that the increase in extreme water levels in the period 1958 - 2005 is larger than the increase in the period 2005 - 2060. Although the impact of the morpho-dynamics on the extreme water levels varies between different areas within the Limfjord, the general picture is clear and significant inasmuch as the natural development has led to an increase in the extreme water levels in the entire Limfjord (perhaps except the easternmost more channel-like parts) since the 1950s.

Table 1: 100 year return heights (cm) from extreme water level statistics

Statistics	Lemvig	Skive	Logstor	Source
1958 modelled and corrected	173	183	199	Christensen (2011a)
2005 modelled and corrected	199	197	203	Christensen (2011a)
2060 modelled and corrected	238	209	220	Christensen (2011a)
2007	183	193	194	Sørensen & Ingvardsen (2007)
2012 corrected	196	195	201	Sørensen *et al.* (2013a)

Selected levels for a statistical 100 year return height at Lemvig, Skive and Logstor in Table 1 show a marked increase between 1958 and today. For reference the calculated statistics from 2007 (Sørensen & Ingvardsen, 2007) and 2012 (Sørensen *et al.*, 2013a), the latter including bathymetry corrected water levels in the statistics, are also shown.
Differences between the 2005 and 2012 statistics are small and are mainly due to the inclusion of all extremes in 2012 and with cut-off levels decided individually at each station. As the statistics have all been related to MSL, SLR must be added to give true number for 2060 in DVR90 and subtracted from 1958 water levels, respectively. At Logstor, the small difference between 1958 and 2005 is mainly due to local off-shore shoals that lead to high surge levels in the town. The largest magnitude of change is found in the Lemvig statistics, where the estimated 100 year return height has increased 13 cm between the last two official DCA statistics (2007 & 2013) and a dramatic increase in storm surge levels is found between the 1958, the 2005 and the 2060 statistics, respectively.

Although many additional morpho-dynamic changes may occur on a local level and modelling has its limitations, too, the results are believed to yield a solid picture of the actual impacts of change in the Limfjord.

4 INCREASED FLOODING HAZARDS DUE TO SLR AND SUBSIDENCE

A study in Thyboron in collaboration between the Danish Geodata Agency, DCA, DTU-Space and the Lemvig Municipality set out to explore uses of the Digital Elevation Model (DEM) in relation to land movement and future storm surges, hereby integrating into coastal climate change adaptation extreme water levels, SLR, glacio-isostatic adjustment (GIA), local subsidence and flooding hazards (Sonne *et al.*, 2012; Vognsen *et al.*, 2013). Being a methodology study preliminary assumptions have been made which are currently being dealt with in relation to qualification of methods, causes, solutions, and with a task to make the method more widely applicable (e.g. Broge *et al.*, 2013).

4.1 Methodology and results

The precision of the DEM (resolution 1.6 m) was evaluated from levelling of manhole covers of the sewer system. Based on 136 of a total of 349 measurements, where the slope gradient was below 0.7° and thus considered representative of the DEM, Figure 9, a deviation of +23 mm (std. dev. = 4 mm) was found and used for adjustment of the DEM.

From precise national levelling campaigns, GPS-measurements and GIA models a glacio-isostatic uplift in the Thyboron area of 1 mm/y is found (Knudsen & Vognsen, 2010; Knudsen *et al.,* 2012) and the current absolute rate of SLR at Thyboron from tide gauge measurements (Knudsen & Sørensen, 2013) is established at 3 mm/y (for use in this study).

Figure 9: General elevation pattern (right) and surface slope map. Manhole covers are either in areas with slope gradients <0.7° (green dots), typically in the middle of roads, or > 0.7° (red)

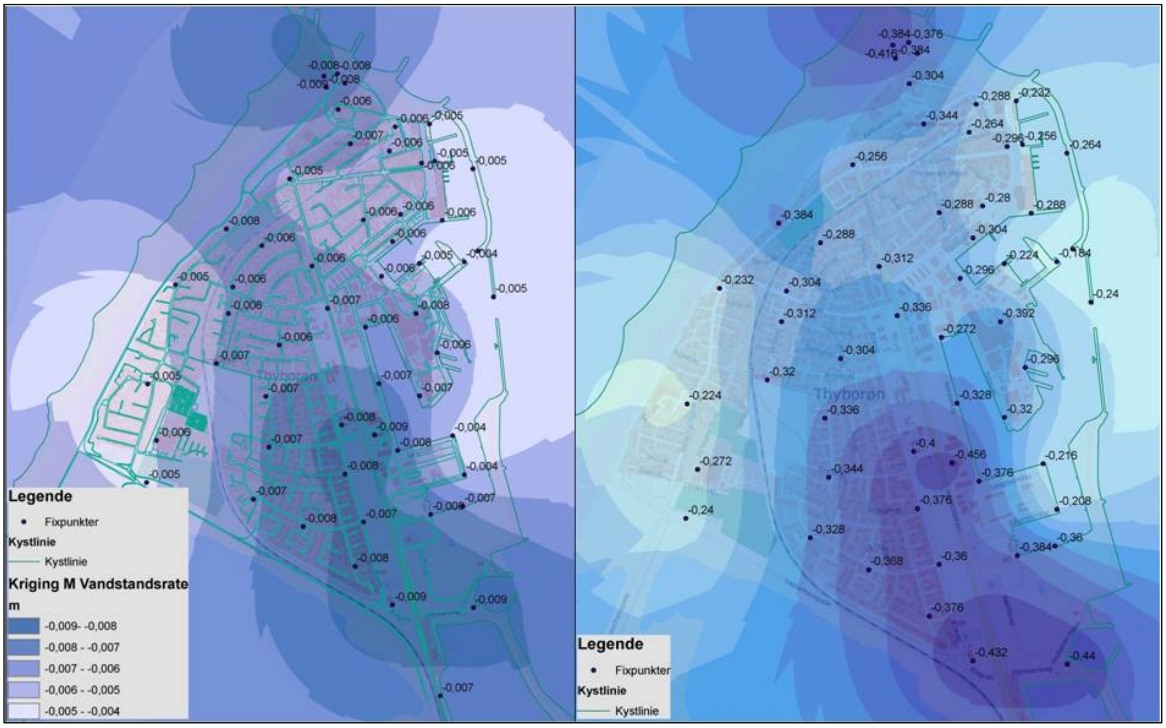


Figure 10: Relative change (m/y) between the land surface and mean sea level today (left) in Thyboron and accumulated change by 2060 provided that rates of change remain constant

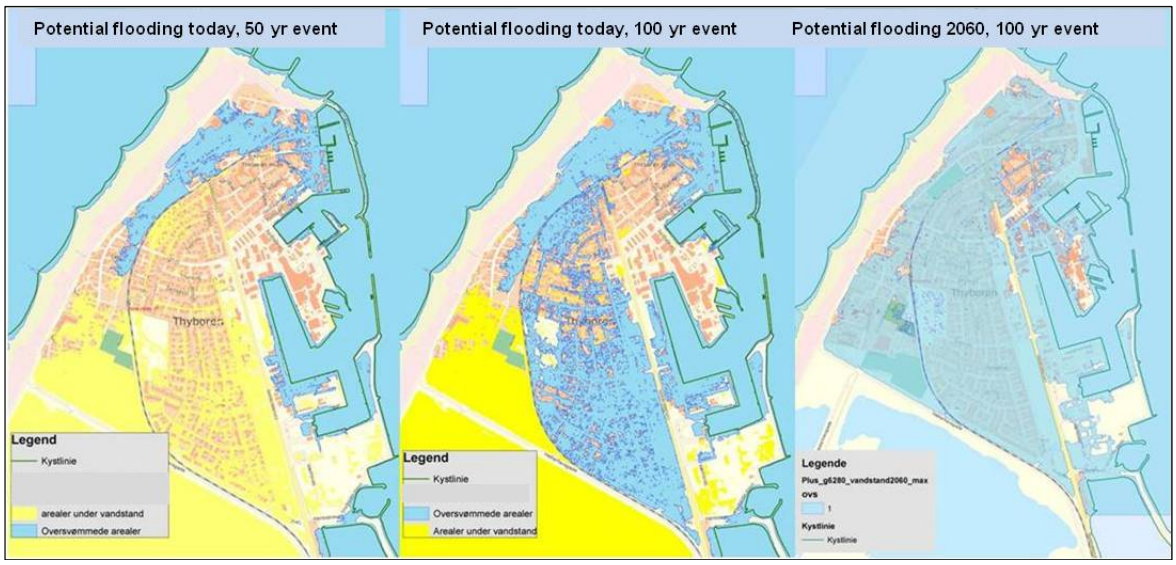


Figure 11: Maximum potential extent of flooding (blue) in Thyboron for 50 yrs (187 cm DVR90) and 100 yrs (193 cm DVR90) return heights today, and a 100 yrs return height in 2060 related to the adjusted DEM (right). Yellow areas are not flooded but lie below maximum water level

Levelling campaigns in 2012, 2009 and 2006 to a closely spaced grid of benchmarks (and inclusion of previous campaigns in 2003, 1998, 1985 and 1970) show a very consistent pattern of local subsidence in Thyboron of 2 - 8 mm/y. This subsidence is largest in areas with landfill but is also governed by the underlying geology. Adding up the individual contributions, the vertical reduction between MSL and land surface in Thyboron is 4 - 10 mm/y as mapped in Figure 10 (left) using a kriging interpolation.

If the changes are projected linearly into the future (assuming constant rates of SLR and land movement) they add up to a maximum of almost half a meter by 2060, Figure 10 (right). This assumption is not strictly valid but is considered sufficient on a 50 year timescale for illustration purposes.

Turning to the extreme water level statistics, figure 11 (left and middle) some flooding today may be anticipated already at a 50 year event and increasing at a 100 year event, when the time factor is not considered. This probably is not the case and may be due to the DEM; however some immediate attention is probably needed in relation to flooding protection. Again, assuming that SLR will not affect the extreme water level statistics, the DEM is made "dynamic" in time and adjusted to the known and projected changes in MSL and land movement to yield a result for the potential flooding extent in 2060, Figure 11 (right). Almost the entire town is then at risk.

The above example thus shows the ability to model SLR and land movement in the DEM to give more realistic future flooding scenarios. It also shows that, in addition to extreme water levels and the potential consequences of SLR, local subsidence may locally be the most important parameter of change at an intermediate timescale. Further, the example emphasizes that attention cannot be given solely to areas in the immediate vicinity of the coastline in coastal climate adaptation as areas further inland may also be susceptible to change over time.

5 DISCUSSION AND CONCLUSIONS

Based on the presented results a simplified conceptual set-up for the location-specific relative change in extreme water levels may be stated as:

$$\Delta WL_{(rel)extreme} = (\Delta MET_{extreme} + \Delta PHYS_{extreme}) + (\Delta SLR_{abs} + \Delta GIA_{abs} + \Delta LOW_{abs}) \quad (1)$$

The first bracket contains effects from a potential increase in future storminess and from physical/ morpho-dynamic change and the second bracket relates to SLR and to land movement due to regional glacio-isostatic adjustments and to local subsidence. The second bracket adds corrections to the de-trended extreme water level statistics for use in assessing the future change but may also affect the water level excursion in extreme events.

The impact in the statistics due to climate change is not dealt with explicitly here, but projections give a hint as to how the future may look and the effects can be modelled or calculated. The influence of morpho-dynamic changes can also be modelled provided that sufficient data is available and one knows what to look for, and local subsidence rates can easily be measured but most often we do not have enough data and knowledge present at hand. The presented studies show, however, that the impact of morpho-dynamics and local subsidence can be significant with contributions by far exceeding the effects of SLR both back in time and in coming decades.

In relation to the Danish statistics both the method applied and the extreme water level data may need a re-evaluation and update, and in relation to storm surge heights it is intriguing to await the future. Nevertheless, Denmark is in the midst of making climate change adaptation plans at the municipality level where robust statistics are "a must" in evaluating current and future flooding hazards. Focus is still often on the uncertainties in projections of SLR and on what to believe in, however.

In relation to the impacts of morpho-dynamics in the Limfjord changes occurred almost unnoticed for decades and local subsidence has not been a major issue there or elsewhere up till now. From a managers point of view this means that projected or believed safety levels against coastal flooding are in many places not met, especially where benchmarks are old and have not been related to a fixed datum for perhaps several decades. Considerations about potential morpho-dynamic changes and local subsidence are therefore advocated in relation to extreme water level statistics, flooding hazards and climate change adaptation in Denmark.

The current work on making a "dynamic" DEM by adjusting the model to future conditions based on (1) seems very promising and may prove a viable and easily accessible tool in relation to flooding issues and e.g. planning and maintenance of sewer systems. Mapping of areas in Denmark potentially susceptible to submergence has been initiated and will, together with the ongoing national flooding hazard mapping based on extreme water level statistics and extreme water level curves, form a solid basis for coastal climate change adaption in Denmark.

Finally, an integration of the two case studies is desirable in relation to the upcoming tasks of determination and implementation of measures to reduce the extreme water levels in the Limfjord area.

6 ACKNOWLEDGMENTS

Co-authors Bo Brahtz Christensen (DHI), Holger Toxvig Madsen (DCA) and Søren Bjerre Knudsen (DCA) are acknowledged for their work on the Limfjord investigations as is Signe M. Ingvardsen (DCA). In relation to the work on land movement and the "dynamic" DEM appreciation primarily goes to Karsten Vognsen, Ian B. Sonne and Niels Broge from the Danish Geodata Agency and to Per Knudsen from DTU-Space for an exciting collaboration containing a multitude of facets that still need to be addressed and investigated.

7 REFERENCES

Broge, N, Sonne, B., Vognsen, V., Knudsen, P. and Sørensen, C. (2013): *Integration af lokale sætninger i klimatilpasning og risikokortlægning*. Dansk Kystkonference 2013, 11. - 12. september 2013, Køge, Danmark.

Christensen, B.B. (2011a): *Stormflodsundersøgelse i Limfjorden. Modelgrundlag, kalibrering og følsomhedsanalyse*. Teknisk Notat. Udarbejdet af DHI for Kystdirektoratet. 265p. + Appendices.

Christensen, B.B. (2011b): *Stormflodsundersøgelse i Limfjorden.– ny bathymetri og uddybning til 10 meters vanddybde*. Teknisk Notat. Udarbejdet af DHI for Kystdirektoratet.

Ingvardsen, S.M., Knudsen, S., Toxvig Madsen, H., Sørensen C. and Bisgaard C. (2012): *Thyborøn Kanal og Vestlige Limfjord*. Teknisk rapport. Kystdirektoratet.

Knudsen, P., Engsager, K., Khan, A., Andersen, O.B., Sørensen, C., Vognsen, K., Sonne, I.B. and Broge, N. (2012): *Landbevægelser i Danmark og deres betydning i forhold til fremtidig havspejlsstigning*. Højdedata - til klimatilpasning og beredskab, Holckenhavn Slot, 2. oktober 2012.

Knudsen, P. and Vognsen, K. (2010): *Metode til at følge vandstandsstigningstakten i de danske farvande*. KMS Technical report no. 08. Danish National Cadastre, Copenhagen.

Knudsen, S.B., Ingvardsen, S.M., Madsen, H.T., Sørensen, C. and Christensen, B.B. (2012): *Increased water levels due to morphodynamic changes; the Limfjord, Denmark*. *Coastal Engineering Proceedings* [Online], ICCE 2012, Santander, Spain, Vol. 1, 33.

Knudsen, S.B., Sørensen, C., Toxvig Madsen, H. and Ingvardsen, S.M. (2011): *Thyborøn Kanal 2009 – Teknisk rapport*. DCA, Lemvig.

Knudsen, S.B. and Sørensen, C. (2013): *Vandstandsmåling og vandstandsudvikling langs Jyllands centrale vestkyst*. Teknisk rapport. Kystdirektoratet.

Sonne, I.B., Broge, N., Vognsen, K., Borg, C., Knudsen, P. and Sørensen, C. (2012): *Dynamisk fremskrivning af Danmarks Højdemodel (DHM) ud fra højde- og vandstandsmålinger i Thyborøn. Højdedata -til klimatilpasning og beredskab*, Holckenhavn Slot, 2. oktober 2012.

Sørensen, C. and Ingvardsen, S.M. (2007): *Højvandsstatistikker 2007*. Kystdirektoratet.

Sørensen, C., Madsen, H.T. and Knudsen, S.B. (2013a): *Højvandsstatistikker 2012*. Kystdirektoratet.

Sørensen, C., Knudsen, S.B., Madsen, H.T., Ingvardsen, S.M. og Christensen, B.B. (2013b): *Thyborøn Kanals indvirkning på stormflodsvandstande i den vestlige Limfjord*. DNO. *17. danske Havforskermøde*, Roskilde, 21.-23. januar 2013.

Vognsen, K., Sonne, N., Broge, N., Sørensen, C. and Knudsen, P. (2013): *Metode til fremskrivning af oversvømmelsesomfang ved stormflod*. Geodatastyrelsen. Technical Report.

Application of a conditional approach for multivariate extreme values to flood risk

Wyncoll, D.[1], Gouldby, B.[1] and Hames, D.[1]

[1]Flood Management, HR Wallingford, Wallingford, UK, Email: d.wyncoll@hrwallingford.com

Abstract

Flood risk is a function of probability and consequence. Flooding events can arise from different sources, pluvial, fluvial, groundwater and coastal flooding. Sometimes flooding at a specific location arises from a single source in isolation. Often however, flooding arises as a combination of two or more flood source variables (e.g. rainfall, river flows, surges and wave heights) interacting. Moreover, flood events can be spatially diverse and it is necessary to model the dependence between extreme events over large spatial areas. Often the variables are partially dependent and assumptions of full dependence or complete independence are difficult to justify.

The application of multivariate extreme value analysis methods, that seek to capture the extreme dependence between different variables, has been the subject of a significant amount of research within the context of flood risk analysis. Examples include extreme, fluvial flows at confluences, estuarine water levels, metrological surges, waves and sea levels. The approaches that have been applied however, have been restrictive in terms of the number of variables that have been considered and the spatial extents. The restrictions have often been imposed by the assumptions within the modelling of dependence between the variables in the upper tail. A multivariate extreme value method has been developed that uses a conditional approach, Heffernan & Tawn (2004). The approach involves transformation of the variables to scales more suited to modelling extreme dependence, prior to fitting a multivariate regression model. There is increasing evidence to suggest that this methodology is capable of providing more robust results to a wide range of environmental variables and flood risk related problems.

This paper describes the underlying methodology, and then demonstrates its application on two case study sites within England. The first case study is a catchment in the North-West of England that historically suffers from fluvial flooding. The second case study focuses on coastal flooding and focuses on a more localized area in the South-West of England.

1 INTRODUCTION

Flood risk is a function of probability and consequence. The probability component includes assessment of the probability of extreme drivers of flooding, heavy rainfall, high river flows, extreme sea conditions, for example. Within the *Source Pathway Receptor* framework of flooding, Sayers *et al.* (2002), these are often referred to as *Source* variables. Although flood events can arise over large spatial scales, the severity of the *Source* variable is, however, unlikely to be constant (equal likelihood or return period) at all spatial locations. There is therefore a need to understand the variability that can arise in terms of the severity of the extreme event at different spatial locations.

Additionally, flooding can arise from multiple variables within one *Source* or from multiple *Sources*. Coastal flooding, for example, can arise from combinations of different surges, wave heights and periods, astronomical tides and wind speeds. Flooding can however, also arise from combinations of high river flows and extreme sea conditions. It is therefore evident that a requirement exists for multivariate extreme value models that are capable of application to a variety of flooding problems.

There are a large number of alternative statistical approaches for modelling multivariate extreme value dependence, for example Coles & Tawn (1991), Joe *et al*. (1992), Coles & Tawn (1994), Ledford & Tawn (1996), Bruun & Tawn (1998) and Nelson (1999). These, and other types of

methods, have been applied to environmental variables that relate to flooding, see Acreman (1994), Svensson & Jones (2002), Hawkes *et al.* (2002), De Michele *et al.* (2007) and Wahl *et al.* (2012), for example. These methods, however, often place restrictive assumptions on the forms of extremal dependence, specifically assuming all pairs of variables are either dependent at extreme levels or fully independent. The Heffernan & Tawn (2004) method removes this constraint by specifically modelling extreme and typical values of one variable conditionally upon extreme values of a second. There is increasing evidence that this approach can be successfully applied to a wide range of environmental variables specifically associated with flood risk, in particular Lamb *et al*. (2010), Jonathan *et al.* (2010), Keef *et al*. (2012), Neal *et al.* (2012) and Wyncoll & Gouldby (2013).

This paper describes the application of the Heffernan & Tawn (2004) multivariate extreme value method to two different flooding problems. The first relates to spatially extensive fluvial flooding in a catchment in the north-west of England. The second relates to coastal flooding from multiple variables on England´s south-west coast.

2 METHODOLOGY

2.1 Multivariate extreme value model

Let $\mathbf{X}=(X_1, \ldots, X_d)$ represent the d source variables that may each occur in extreme as well as typical values. The problem may depend upon other source variables that do not occur in extreme values, such as wind direction, but these are assumed to be dependent upon the variables in **X**, at least at extreme levels. A dataset of independent joint events of all source variables is required where a peak value is given for each extreme variable in an event. For simplicity of presentation, we assume that flooding is related to extremely large rather than extremely small values of each of the variables X_i. Where this is not the case, the extreme value analysis may be conducted using the negated variable $-X_i$ taking care to negate the sampled values at the end of the analysis.

The first step in the Heffernan & Tawn (2004) method is to model extreme values of each of the d variables marginally. For this, the standard peaks-over-threshold approach of Davison and Smith (1990) is applied. Independent peak values above a high threshold u_i are fitted to the generalized Pareto distribution (GPD) which takes the form

$$P(X_i>x \mid X_i> u_i)= \left[1+\xi_i \frac{(x- u_i)}{\beta_i}\right]_+^{-1/\xi_i} \qquad \text{for } x> u_i \tag{1}$$

where $\beta_i>0$ and $\xi_i \in R$ are parameters and $[y]_+=\max(y, 0)$. The threshold u_i is chosen to be just large enough to ensure a stable estimate for the shape parameter ξ_i for all larger thresholds.

The GPD model above a threshold is combined with the empirical distribution $\widetilde{F}_i$ of the dataset below the threshold to give the following semi-parametric estimate of each cumulative distribution function, first used by Coles & Tawn (1991):

$$\widehat{F}_i(x)= \begin{cases} \widetilde{F}_i(x) & \text{for } x \le u_i \\ 1- \left(1- \widetilde{F}_i(x)\right) \left[1+\xi_i \frac{(x- u_i)}{\beta_i}\right]_+^{-1/\xi_i} & \text{for } x> u_i \end{cases} \tag{2}$$

This is used to standardize the d source variables to common Gumbel marginal distributions via the probability integral transform to give

$$Y_i= -\log\left(-\log\left(\widehat{F}_i(X_i)\right)\right). \tag{3}$$

The transformed variables $\mathbf{Y}=(Y_1, \ldots, Y_d)$ retain the dependence structure of the original dataset but not the marginal characteristics. This transformation to common margins is often known as the copula approach.

Let $\boldsymbol{Y}_{-i}$ denote the vector of all variables Y_j excluding Y_i. The Heffernan and Tawn (2004) approach consists of modelling the variables $\mathbf{Y}_{-i}$ conditionally upon Y_i being extreme. This is then repeated for all extreme variables i=1, …, d. For fixed i, the vector $\mathbf{Y}_{-i}$ is typically modelled using the multivariate non-linear regression model

$$\mathbf{Y}_{-i} \mid Y_i = \mathbf{a}\, Y_i + Y_i^{\mathbf{b}}\, \mathbf{Z} \qquad \text{for } Y_i > v \tag{4}$$

where v is a high threshold, $\mathbf{a} \in [0, 1]$ and $\mathbf{b}<1$ are vectors of parameters and $\mathbf{Z}$ is a vector of residuals. Vector arithmetic should be interpreted component-wise so that each Y_j is modelled as a function of Y_i using parameters $a_{j|i}$ and $b_{j|i}$ and residual $Z_{j|i}$.

The regression parameters $a_{j|i}$ and $b_{j|i}$ are estimated using maximum likelihood under the temporary assumption that $Z_{j|i}$ follows a Normal distribution with unknown mean and variance. This fit uses all pairs (Y_i, Y_j) corresponding to cluster maxima of $Y_i>v$ to be consistent with the marginal GPD fits made to cluster maxima of X_i. Asymptotically, $Y_i>v$ is statistically independent of the residual $Z_{j|i}$ hence the threshold v is chosen to be just large enough for this condition to hold. Once all parameter estimates have been found a non-parametric estimate of the joint distribution of $\mathbf{Z}$ is constructed from the empirical distribution of the sample residuals. Any additional source variables, such as wind direction, are also bound to this empirical distribution so that every sample of $\mathbf{Z}$ has a corresponding value of all such non-extreme variables.

The above description assumes variables Y_i and Y_j occur concurrently so is not appropriate for modelling temporally dependent data with a time lag between peak values. Keef *et al.* (2009) overcome this deficiency by fitting the conditional model of $Y_{j,\, t+\tau} \mid Y_{i,\, t}$ for a selection of lags τ for each variable $j \neq i$. This allows the model to be used to simulate new events over a range of lags so that the largest values in each time window need not occur concurrently.

2.2 Simulation of extreme events

The fitted statistical model provides parameter estimates $a_{j|i}$ and $b_{j|i}$ for all pairs $i \neq j$ in addition to empirical samples of joint residuals $\mathbf{Z}$ for each source variable i. If the extension of Keef *et al.* (2009) for temporally dependent data is used, a separate set of parameters and residuals is produced for each of a set of lags τ. These constitute the fitted statistical model and can be used to simulate an arbitrarily large number of dependent events, each of which is extreme in one or more of the original source variables. Such simulated events may be used in many ways, depending on context (see the case studies in Section 3).

To simulate a single event, an extreme variable i must be first selected. An extreme value for this variable, Y_i, is sampled above a high threshold v and the remaining variables are sampled from the fitted model for $\mathbf{Y}_{-i} \mid Y_i$ conditional upon being less extreme than Y_i. Amongst a large sample of extreme events, each variable is selected to be most extreme for a proportion matching that observed in the fitted dataset. The sample of extreme events is finally transformed from the Gumbel scale to the original scale by inverting the probability integral transform applied to the fitted data.

The simulation consists of repeating the following steps after selecting an extreme variable i:

1. Sample a value Y_i from the standard Gumbel distribution conditioned to exceed v.
2. Independently select one of the joint residuals $\mathbf{Z}$ for variable i.
3. Calculate $Y_j = a_{j|i}\, Y_i + Y_i^{b_{j|i}}\, Z_{j|i}$ for all $j \neq i$.
4. Reject the joint sample $\mathbf{Y}$ unless Y_i is maximum over all source variables.
5. If accepted, set $X_j = \hat{F}^{-1}\left(\exp\left(-\exp\left(-Y_j\right)\right)\right)$ for all j to transform the sample back to the original scale.

If the temporally dependent extension is used, the sample Y_j is maximum over all lags τ, see Keef *et al.* (2009). If there are additional non-extreme source variables, these are resampled from the dataset by selecting the values corresponding to the joint residual $\mathbf{Z}$ selected for each sample. An example of this is given in the second case study in Section 3.2.

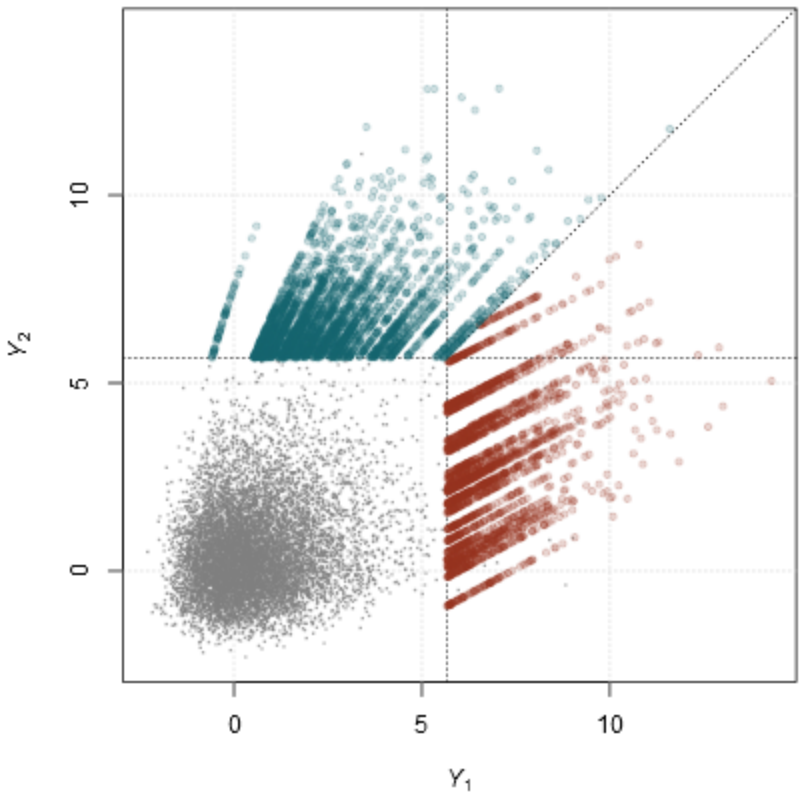


Figure 1: Example Heffernan and Tawn samples (red and blue) plotted on the Gumbel scale against the observed data

Example simulated events in two dimensions are shown in Figure 1 plotted on the Gumbel scale. The red points were created by sampling Y_1 above the threshold v = 5.6 and setting $Y_2 \mid Y_1$ from the fitted regression equation for a randomly selected residual. Each line of points corresponds to a set of samples all generated from the same residual. The point is rejected unless Y_1 is largest, i.e. more extreme than Y_2 given the common margins. The blue points are similarly samples for which Y_2 is most extreme. Between them, they provide 100 times as many extreme values, taking values above v to be extreme, than were observed in the raw dataset.

3 CASE STUDIES

3.1 Application to fluvial extremes

This first case study uses the method of Heffernan & Tawn (2004) with the extensions of Keef *et al*. (2009) to model the temporal-spatial dependence of extreme river flows in the Eden catchment. The Eden catchment in the north-east of England covers approximately 2,400 km^2. In the higher regions to the south, the average annual rainfall exceeds 2,800 mm and is over three times the national average of England and Wales (920mm), Environment Agency (2009a). The main population centres are Carlisle, Penrith and Appleby. Carlisle has a history of flooding with significant events having been recorded in the past and most recently in January 2005, when nearly 3,000 homes were flooded, three people died and the resulting flood losses were estimated as £400m, Geographical Association (2009).

In this case study, the source variables X_i were peak flows at d = 248 representative sites spanning the river network. Concurrent time series data spanning 27 years was provided at each of these sites from a gridded hydrological model (Grid-to-Grid, or "G2G", of Bell *et al.*, 2007) interpolated to each of the sites. Peak flow clusters were first identified for each site using the runs method of Smith & Weissman (1994) and the cluster maxima above high thresholds were fitted to GPD. Each peak value was paired to local maxima at all other sites to identify a total of 709 joint events, each of which was extreme for at least one of the sites.

After transformation to Gumbel scales, the lagged Heffernan and Tawn regression model was fitted using each site as a conditioning variable and for each of a range of lags up to days. From this, years' worth of extreme river flows were simulated from the fitted model, using rejection sampling to control the proportion of events that were most extreme at each site. Each sample comprised a peak flow level at all sites with at least one extreme flow. The time lags between the sample peaks at each site were also available as resampled when the residual was selected.

The simulated events for five selected sites are plotted in Figure 2 as paired scatter plots. The diagonal indicates the location of each site on a map of the catchment area. This demonstrates the range of dependencies the Heffernan and Tawn model is able to represent. Unsurprisingly, the sites in close proximity display a strong dependence while sites far apart are only partially dependent upon each other.

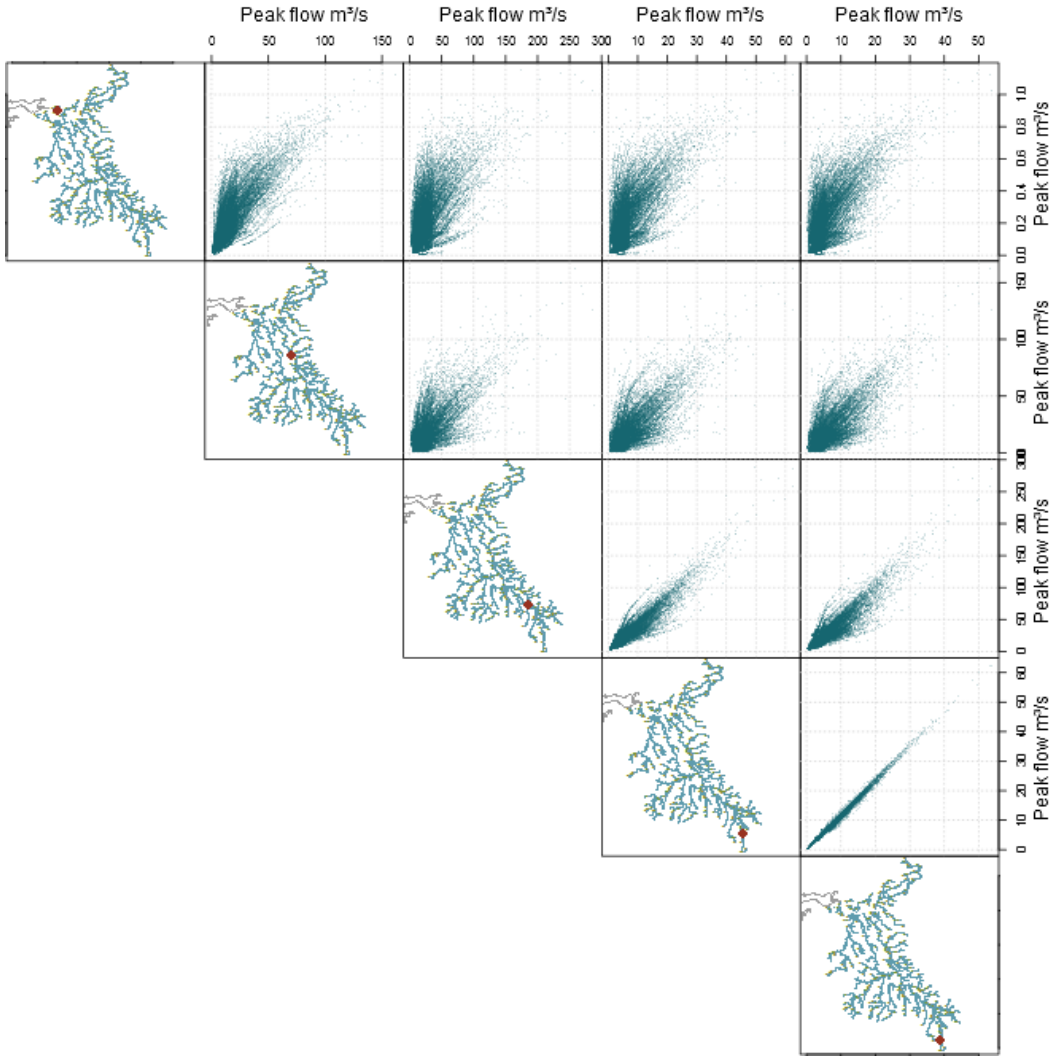


Figure 2: Scatter plots of multivariate extreme samples for every pair of five sites in the fluvial case study

Having generated a large set of simulated extreme events these were then used as input to a comprehensive flood risk analysis model. The risk analysis model considers the reliability of the flood defences as well as inundation and economic damage components, Gouldby *et al.* (2008). The model is currently used by the Environment Agency in England to assess the national flood risk, Environment Agency (2009b). The method currently applied in practice involves incorporating a simplifying assumption of full dependence between the extreme flow events. The extreme event set generated using the multivariate statistical model was used to form the boundary conditions to the risk analysis model. It was demonstrated that this approach was able to overcome the simplified method used within the current system, see Wyncoll & Gouldby (2013).

3.2 Application to coastal extremes

This second case study uses the Heffernan & Tawn (2004) method in combination with the SWAN nearshore wave transformation model, Booij *et al.* (1999), and the BAYONET wave overtopping tool, Kingston *et al.* (2008), to assess extreme peak wave overtopping rates at six locations covering a length of coastline of approximately 13 km. There are a range of defence types in this area, including soft defences around Dawlish Warren and a variety of hard defences along the coastal railway line and the towns of Teignmouth, Torquay and Paignton.

Peak overtopping rates for these defences have been assessed for six source parameters: resultant significant wave height H_s, steepness s, resultant wave direction θ, sea level SL, wind speed U and wind direction θ_U. Steepness is defined as a function of significant wave height and resultant wave period T_m by

$$s = \frac{2\pi H_s}{g T_m^2}. \tag{5}$$

Steepness is used in place of wave period as this is likely to be approximately constant for extreme wave heights and can therefore be considered as a non-extreme source variable, Hawkes *et al.* (2002). As directions cannot occur in extreme values, this leaves only $d = 3$ potentially extreme source variables: wave height H_s, sea level SL and wind speed U.

Sea level data since 1991 is available from the National Tidal and Sea Level Facility (www.ntslf.org) for Weymouth to the east and Devonport to the west. Predicted wave and wind data since 1986 is available from the UK Met. Office. Accounting for time lags at the locations of interest and the different data sets as well as missing data, these were amalgamated to give just over 14 years of concurrent data at high sea level conditions for all six source variables.

The standard Heffernan & Tawn (2004) method was applied to $\mathbf{X} = (H_s, SL, U)$ with s, θ and θ_U sampled from their respective empirical distributions. The GPD was first fitted to each of the extreme variables which were then transformed to Gumbel scales. The Heffernan and Tawn regression model was fitted to every pair of Gumbel variables. From the fitted distributions, 20,000 years' worth of extreme events were simulated using rejection sampling to control the proportion of events that are most extreme for each of the three primary variables.

Each realisation comprises a value for each of the six original source parameters and is extreme for at least one of the three primary variables. The samples are shown as paired scatter plots in Figure 3. The red samples are most extreme for wave height H_s, green for sea level SL, blue for wind speed U and the observed data are shown in grey. Between them, they represent the extrapolated joint distribution with an appropriate dependence structure.

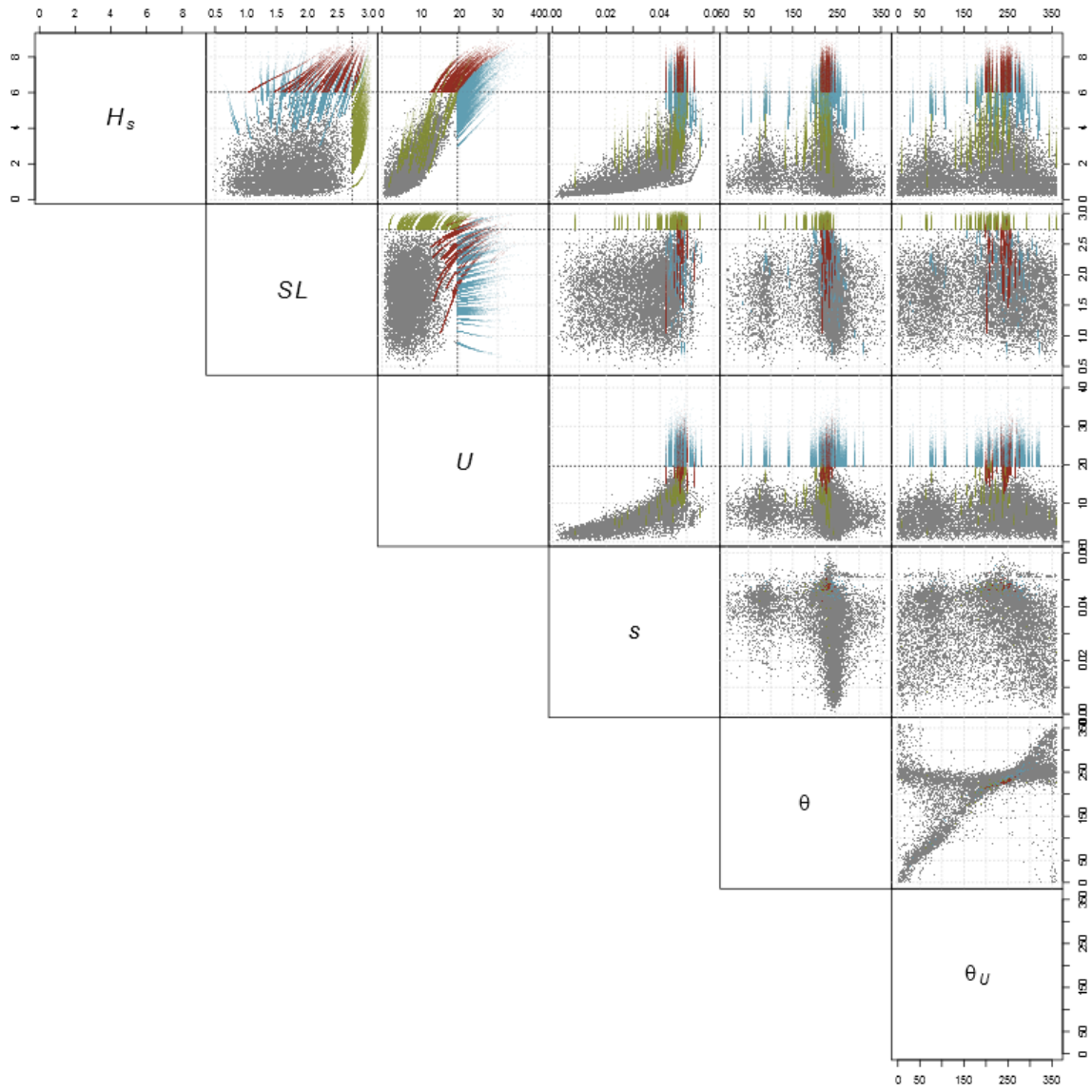


Figure 3: Scatter plots of multivariate extreme samples for every pair of the six source parameters in the coastal case study

Having generated this large set of simulated results, these were then transformed to three nearshore points just seaward of the breaker zone using the SWAN wave transformation model, Booij *et al.* (1999), accounting for all six source variables. These data were then transformed to the structure toe of each defence using a profile model that accounts for wave refraction, shoaling and breaking. For each combination of wave height, period, direction and sea level, the BAYONET wave overtopping tool was then used to estimate peak overtopping rates for each high sea level condition. Using the peak overtopping series for all six locations, empirical distributions of overtopping rates, that included extreme values, were then determined for each defence considered.

Comparisons with overtopping rates calculated for the UK National Flood Risk Assessment (NaFRA) indicated noticeable variations, with the NaFRA results tending to be over-estimated particularly for low return period events, see Figure 4. The reasons for these differences include the lack of spatial resolution within NaFRA and lack of nearshore wave transformation processes. These differences are discussed further in Environment Agency (2013).

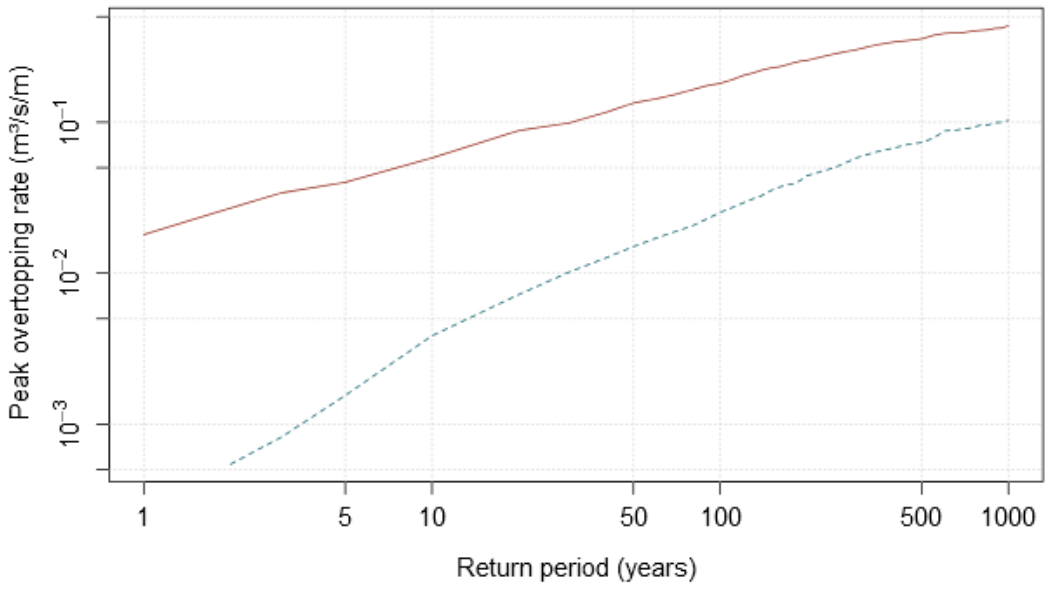


Figure 4: Comparison between overtopping rates from the NaFRA database (red) and the overtopping model (blue dashed) for a defence at Torquay, after Environment Agency (2013)

4 CONCLUSIONS

Flood risk is a function of probability and consequence. Flood risk analysis often requires consideration of extreme events occurring over large spatial scales and/or from multiple *Source* variables. Multivariate extreme value methods are therefore frequently applied yet more traditional multivariate extreme value methods often contain limitations relating to the dependence structure that can limit the spatial scale of the analysis or the number of variables that can be considered. The multivariate method of Heffernan & Tawn (2004) overcomes a number of the limitations of previous methods in relation to the dependence structure within the extremes. The flexibility of the method in its application to flood risk has been demonstrated here on two cases study sites. The first application relates to fluvial flood risk over a relatively large spatial area. The second application, to coastal flooding, shows how the method can be extended to multiple types of variable.

5 ACKNOWLEDGMENTS

The authors are grateful to Environment Agency Staff, David Hornby and Tim Hunt, as well as colleagues, Dr Peter Hawkes, Dr Tim Pullen and Mr Nigel Tozer for discussions relating to the development of the methodology for the coastal analysis.

6 REFERENCES

Acreman, M. C. (1994): *Assessing the Joint Probability of Fluvial and Tidal Floods in the River Roding.* Water and Environment Journal 8 (5), pp.490–496.

Bell, V., Kay A., Jones R. and Moore R. (2007): *Development of a high resolution grid-based river flow model for use with regional climate model output*. Hydrology and Earth System Sciences 11 (1), pp.532–549.

Booij, N., Ris, R. C. and Holthuijsen, L. H. (1999): *A third-generation wave model for coastal regions: 1. Model description and validation*. Journal of Geophysical Research: Oceans 104 (C4), pp.7649–7666.

Bruun, J. T. and Tawn, J. A. (1998): *Comparison of approaches for estimating the probability of coastal flooding.* Journal of the Royal Statistical Society, Series C (Applied Statistics) 47 (3), pp.405–423.

Coles, S. G. and Tawn, J. A. (1991): *Modelling Extreme Multivariate Events*. Journal of the Royal Statistical Society, Series B (Statistical Methodology) 53 (2), pp.377–392.

Coles, S. G. and Tawn, J. A. (1994): *Statistical Methods for Multivariate Extremes: An Application to Structural Design.* Journal of the Royal Statistical Society, Series C (Applied Statistics) 43 (1), pp.1–48.

Davison, A. C. and Smith, R. L. (1990): *Models for exceedances over high thresholds.* Journal of the Royal Statistical Society, Series B (Statistical Methodology) 52 (3), pp.393–442.

De Michele, C., Salvadori, G., Passoni, G. and Vezzoli, R. (2007): *A multivariate model of sea storms using copula*s. Coastal Engineering 54 (10), pp.734–751.

Environment Agency (2009a): *Eden Catchment Flood Management Plan: Summary Report 2009* E. Agency.

Environment Agency (2009b): *Flooding in England: A National Assessment of Flood Risk.*

Environment Agency (2013): *Prototype wave overtopping tool for MDSF2 B*. Environment Agency Bristol.

Geographical Association (2009): *Managing flood risk - Carlisle Case Study*. http://www.geography.org.uk/resources/flooding/carlisle.

Gouldby, B., Sayers, P. and Tarrant, O. (2008): *Application of a flood risk model to the Thames Estuary for economic benefit assessment Risk Analysis VI: Simulation and Hazard Mitigation*, Caephalonia, WITpress.

Hawkes, P. J., Gouldby, B. P., Tawn, J. A. and Owen, M. W. (2002): *The joint probability of waves and water levels in coastal engineering design*. Journal of Hydraulic Research 40 (3), pp.241–251.

Heffernan, J. E. and Tawn, J. A. (2004): *A conditional approach for multivariate extreme values (with discussion).* Journal of the Royal Statistical Society, Series B (Statistical Methodology) 66 (3), pp.497–546.

Joe, H., Smith, R. L. and Weissman, I. (1992): *Bivariate Threshold Methods for Extremes*. Journal of the Royal Statistical Society, Series B (Statistical Methodology) 54 (1), pp.171–183.

Jonathan, P., Flynn, J. and Ewans, K. (2010): *Joint modelling of wave spectral parameters for extreme sea states.* Ocean Engineering 37(11–12): 1070–1080.

Keef, C., Tawn J. and Svensson C. (2009): *Spatial risk assessment for extreme river flows*. Applied Statistics 58 (5), pp.601–618.

Keef, C., Tawn, J. and Lamb, R. (2012): *Estimating the probability of widespread flood events.* Environmetrics 24 (1),pp.13–21.

Kingston, G., Robinson, D., Gouldby, B. and Pullen, T. (2008): *Reliable prediction of wave overtopping volumes using Bayesian neural networks. FLOODrisk 2008,* Keble College, Oxford, UK.

Lamb, R., Keef, C., Tawn, J., Laeger, S., Meadowcroft, I., Surendran, S., Dunning, P. and Batstone, C. (2010): *A new method to assess the risk of local and widespread flooding on rivers and coasts*. Journal of Flood Risk Management 3 (4), pp.323–336.

Ledford, A. W. and Tawn, J. A. (1996): *Statistics for near independence in multivariate extreme values*. Biometrika 83 (1), pp.169–187.

Neal, J., Keef, C., Bates, P., Beven, K. and Leedal, D. (2012): *Probabilistic flood risk mapping including spatial dependence.* Hydrological Processes 27 (9), pp.1349–1363.

Nelson, R. B. (1999): *An Introduction to Copulas*. New York., Springer-Verlag.

Sayers, P., Hall, J. and Meadowcroft, I. (2002): *Towards risk-based flood hazard management in the UK*. Civil Engineering 150, pp.36–42.

Smith, R. L. and Weissman, I. (1994): *Estimating the extremal index.* Journal of the Royal Statistical Society, Series B (Statistical Methodology) 56 (3), pp.515–528.

Svensson, C. and Jones, D. A. (2002): *Dependence between extreme sea surge, river flow and precipitation in eastern Britain*. International Journal of Climatology 22 (10), pp.1149–1168.

Wahl, T., Mudersbach, C. and 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 (2), pp.495–510.

Wyncoll, D. and Gouldby, B. (2013): *Application of a multivariate extreme value approach to system flood risk analysis.* Journal of flood risk management accepted.

Assessing Bivariate Hydrological Design Parameters Under Nonstationary Conditions

Jens Bender[1,2], Thomas Wahl[1,3] and Jürgen Jensen[1,2]

[1]Institute of Advanced Studies – FoKoS, University of Siegen, Germany, Email: jens.bender@uni-siegen.de

[2]Research Institute for Water and Environment, University of Siegen, Germany

[3]College of Marine Science, University of South Florida, USA

Abstract

In the past years the application of copula functions for multidimensional modelling of hydrological parameters has increased. The main advantage of using copulas over other multivariate statistical methods is that the dependence structure of the variables can be modelled separately from their marginal distributions. It is therefore not surprising that since 2010 more than 100 papers have been published related to copula functions in hydrology. However, to the author's knowledge all of the approaches assume stationarity in the marginal distribution parameters as well as in the dependence structure of the variables. This is because the available time series are often too short for using a non-stationary multivariate model. In this study we analyse the joint probability of flood peak and volume based on a 191 years long discharge time series of the Rhine River. In a first step we estimate the time depending marginal distribution parameters (with the L-moment method) of the Generalized Extreme Value distribution with a 50-year moving time window. We find significant positive trends in the location parameter as well as in the shape parameter of both marginals. The significance of the trends is tested using the Mann-Kendall test on the 95% confidence level. Next, we fit Archimedean copula functions, namely the Gumbel, Clayton, Frank and Ali-Mikhail-Haq copulas, to the pseudo-observations of the moving window series. The goodness of the fits is tested using a parametric bootstrapping procedure. For most of the time steps the Frank copula fits best, which implies that there are no major changes in the tail dependence structure of the two variables. However, we find a significant negative trend of the copula parameter θ over the time. Furthermore the variability of θ ranges from 3.0 up to 7.8, which corresponds to a variability of Kendall's τ between 0.31 and 0.59. The influence of the nonstationary behaviour of the variables is illustrated by calculating the joint probability of the flood peak and volume for four cases: i. considering all parameters as time dependent, i.e. the location, scale and shape parameter of the marginals and the copula parameter, ii. considering the location and scale parameter of the marginals and the copula parameter as time dependent, iii. considering the location parameter of the marginals and the copula parameter as time dependent, and iv. considering only the copula parameter as time dependent. The results highlight that the joint probability, illustrated by the quantile-isoline, varies significantly over the time.

1 INTRODUCTION

Over the last years copula functions have been used for several multivariate hydrological analyses. They were applied for rainfall frequency analysis (e.g. De Michele & Salvadori, 2003; Grimaldi & Serinaldi, 2006; Zhang & Singh, 2007), flood frequency analysis considering peak flow and flood volume (e.g. Favre *et al.*, 2004; Zhang & Singh, 2006; Karmakar & Simonovic, 2009), drought frequency analysis (e.g. Shiau, 2006; Kao & Govindaraju; 2010; Song & Singh, 2010a, 2010b), storm surge modeling (e.g. Wahl *et al.*, 2012), and for several other multivariate problems. The main advantage of using copulas over other multivariate statistical methods, and one of the reasons for the numerous applications, is that the dependence structure of the variables can be modeled separately from their marginal distributions. Furthermore copulas are relatively easy to construct and capable of modeling a broad range of dependence structures. To the author's knowledge all previous hydrological studies using copulas assumed stationarity in the marginal distributions as well as in the dependence structure. This assumption, however, might not be correct as nonstationary

behavior of one or more marginal parameters and/or the dependence can influence the results of a multivariate statistical analysis significantly. In the univariate case, several studies have been carried out exploring nonstationary extreme value models. Mudersbach & Jensen (2010), for example, derived future coastal design water levels using a nonstationary approach.
In case of multivariate trends, i.e. a time depending change in the dependence structure between 2 or more variables, no studies have been carried out so far. Chebana *et al.* (2013) already suggested considering copula functions with time dependent parameters. There are several reasons that no studies concerning multivariate, nonstationary statistics are available. From a practical point of view, one main reason is that time series are often too short to set up a nonstationary multivariate model. On the other hand there exist still some mathematical problems which are not solved yet, e.g. how to handle the problem that the marginal distribution family varies over time.
In this study we combine the univariate nonstationary extreme value analysis with a novel approach to model changes in the dependence of two variables over the time. The aim is to show how bivariate design parameters may change if the nonstationary behavior of the system is taken into account. We use a 192 years long time series of mean daily discharge of the Rhine River at gauge Worms and model the annual maxima of the flood discharge and the corresponding flood volumes.
In section 2 we give a short introduction to the copula theory and to univariate and bivariate nonstationary extreme value analysis. In section 3 the data set is introduced, before the results are presented in section 4 and discussed in section 5.

2 METHODS

2.1 General Copula Theory

Since copula functions have been used during the last years in hydrology and a corresponding number of papers have been published only a short introduction to the theoretical background of copula functions is given here. More information can be found for example in Nelsen (2006), who provided a detailed introduction to the subject.

Copulas are flexible joint distributions for modeling the dependence structure of two or even more random variables. First mentioned by Sklar (1959), the joint behavior of two (or more) random variables X and Y with continuous marginal distributions $u = F_X(x) = P(X \leq x)$ and $v = F_Y(y) = P(Y \leq y)$ can be described uniquely by an associated dependence function or copula function C. In the bivariate case, the relationship between all $(u,v) \in [0,1]^2$ can be written as

$$F_{X,Y}(x,y)=C[F_X(x),F_Y(y)]=C(u,v) \quad (1)$$

where $F_{X,Y}(x,y)$ is the joint cumulative distribution function (cdf) of the random variables X and Y.
A copula function with a strictly monotonically decreasing generator function $\varphi: [0,1] \rightarrow [0,\infty]$ with $\varphi(1) = 0$ belongs to the Archimedean copula family. The general form of one-parametric Archimedean copulas is

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

where θ denotes the copula parameter. In this study we use four Archimedean copulas, namely the Clayton, Frank, Gumbel, and Ali-Mikhail-Haq copulas. They are relatively easy to construct, flexible and capable of modeling the full range of tail dependencies. The Clayton copula has lower tail dependence, while the Frank copula has no tail dependence and the Gumbel copula has strong upper tail dependence (Schölzl & Friedrichs, 2008). The Ali-Mikhail-Haq copula is also considered in case of a weak correlated dependence structure. The goodness of the fit is tested by employing the parametric bootstrapping procedure proposed by Genest *et al.* (2009), which compares the so called "empirical copula" with the parametric copula using the Cramér–von Mises statistics S_n. Large S_n values lead to the rejection of the null hypothesis that the present (unknown) copula C of the sample belongs to the chosen parametric copula family.

2.2 Non-stationary Extreme Value Statistic

In a multivariate framework, nonstationarity can emerge in the statistical attributes of the univariate variables, in the dependence structure of the variables, or both. To capture the possible non-stationarity of the marginals, we fit the time dependent Generalised Extreme Value distribution (GEV) to the annual maxima flood peaks and volumes. As mentioned in Coles (2001) the nonstationary GEV can be written as

$$GEV(x,t) = \exp\left[-\left(1+k(t)\cdot\frac{x-\mu(t)}{\sigma(t)}\right)^{-\frac{1}{k(t)}}\right] \quad (3)$$

where x denotes the variable, μ(t) the time dependent location parameter, σ(t) the time dependent scale parameter and k(t) the time dependent shape parameter.
In order to capture the nonstationary dependence, the well known Archimedean copula functions need to be extended by adding an additional time parameter t. Table 1 summarizes the extended copula functions with the time dependent parameter θ(t).

Table 1: Considered copula functions with time dependent copula parameter θ(t), time dependent generator function φ(k,t), the range of the copula parameter θ and the functional relationship between θ(t) and Kendall's τ(t)

Copula function $C_{\theta(t)}$	Generator φ(k,t)**	Range of θ	Functional relationship of θ(t) to τ(t)
Clayton or Cook-Johnson $\left[u^{-\theta(t)}+v^{-\theta(t)}-1\right]^{\frac{-1}{\theta(t)}}$	$k^{-\theta(t)}-1$	$[0,\infty)$	$\frac{\theta(t)}{\theta(t)+2}$
Frank $-\frac{1}{\theta(t)}\ln\left[1+\frac{\left(e^{-\theta(t)u}-1\right)\left(e^{-\theta(t)v}-1\right.}{e^{-\theta(t)}-1}$	$-\ln\left(\frac{\left(e^{-\theta(t)k}-1\right)}{e^{-\theta(t)}-1}\right)$	$-\ln(-\infty,\infty)\backslash\{0$	$1-\frac{4}{\theta(t)}[1-D_1(\theta(t))]^*$
Gumbel or Gumbel-Hougaard $\exp\left\{-\left[(-\ln u)^{\theta(t)}+(-\ln v)^{\theta(t)}\right]^{\overline{\theta(}}\right.$	$(-\ln k)^{\theta(t)}$	$[1,\infty)$	$1-\theta(t)^{-1}$
Ali-Mikhail-Haq $\frac{uv}{1-\theta(t)(1-u)(1-v)}$	$\ln\left(\frac{1+\theta(t)(k-1)}{k}\right.$	$[-1,1)$	$\frac{3\theta(t)-2}{3\theta(t)}-\frac{2(1-\theta(t))^2\ln(1-\theta(t}{3\theta(t)^2}$

* 1. Debye Function: $D_1(\theta(t))=\frac{1}{\theta(t)}\int_0^{\theta(t)}\frac{k}{e^k-1}dk$

** k = u or k = v

3 DATA

This study is meant to be exemplarily and carried out using the (design-)relevant parameters flood peak (Q) and flood volume (V) at the Rhine River. At gauge Worms discharge measurements have been recorded regularly since 1819 without any major discontinuities (bfg, 2012). The catchment area is 68,827 km². The records used are daily mean discharge data from 1 November 1820 to 31 October 2011, resulting in a time series of 191 years of length. The time series contains no gaps and was checked for homogeneity by the WSA Mannheim. The overall mean of the daily mean discharge from 1820 to 2011 amounts to MQ = 1382 m³/s with a standard deviation of s = 583 m³/s. The highest discharge in the time series was observed on 29 December 1882 with HHQ = 5498 m³/s (see Figure 1).
In a first step we fit the GEV to a moving time window series using the L-moments parameter estimation as proposed by Hosking & Wallis (1997). The time window length is chosen to n = 50 years. For this time window length a stationary behavior is assumed (for fitting the marginal distributions). Furthermore 50 years provide a sufficient number of values for fitting the bivariate

distribution functions (copulas). For the annual maximum discharges, the results of the time dependent distribution parameters μ (location), σ (scale) and k (shape) are shown in Figure 1. Although Coles (2001) assumed that the uncertainties increase unproportionally when considering the shape parameter k as time dependent, for the sake of completeness we consider this parameter as nonstationary, too. Linear trend estimates are performed for all three distribution parameters. The significance of the trend is assessed using the Mann-Kendall test (Mann, 1945; Kendall, 1970) considering the 95% confidence level. The location and scale parameters have positive linear trends as illustrated by the red dashed lines in the middle panels of Fig. 1. The trends for μ and σ amount to $s_{\mu,Q}$ = 2.21±0.15 m³/s/a and $s_{\mu,Q}$ = 2.54±0.17 m³/s/a, respectively The shape parameter k shows a slight but significant negative trend of $s_{k,Q} = -0.002 \pm 1.13 \cdot 10^{-4}$ /a (lower panel of Fig. 1). The black solid lines represent the parameters for the homogenized, i.e. de-trended, data set.

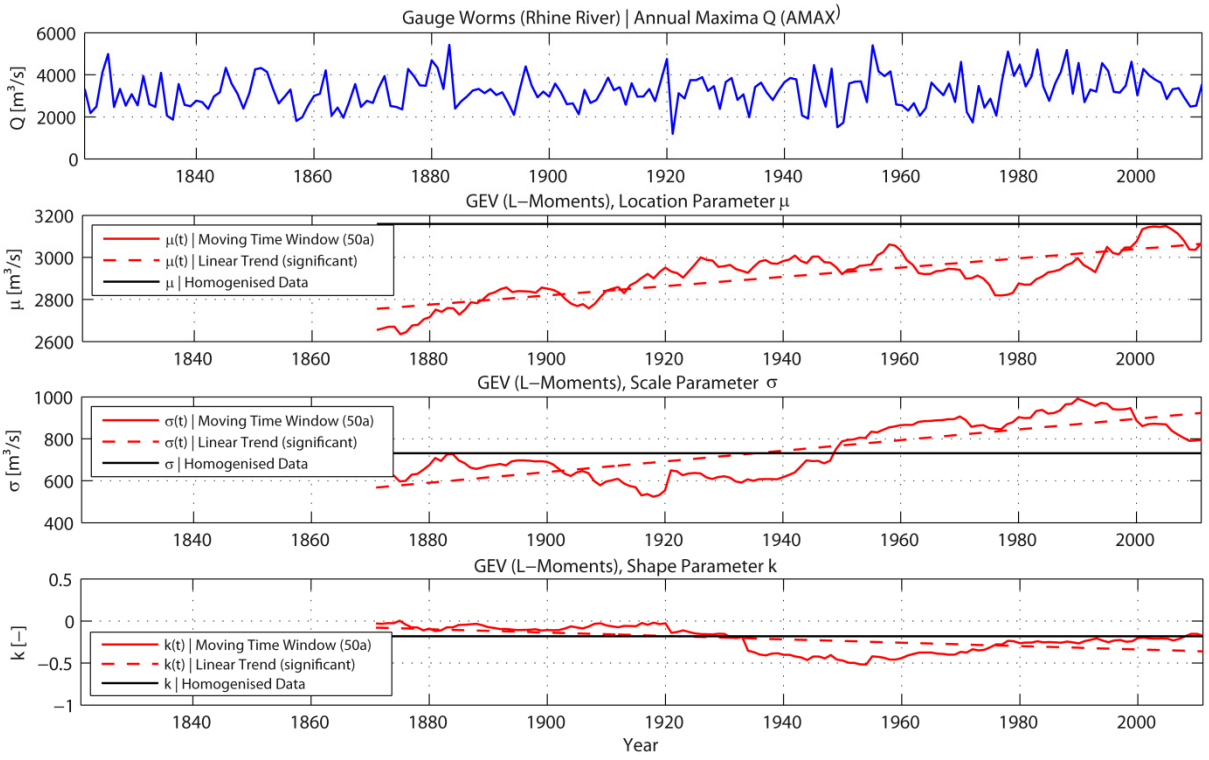


Figure 1: Time series of annual maximum discharges (AMAX) and the time dependent series of the location, scale and shape parameters of the GEV. The parameters are estimated with the L-moments method using a 50-years moving time window. The dashed red lines represent the linear trends and the solid black lines the parameter values of the homogenized data

Equivalent to the time dependent distribution parameters of the AMAX discharges we fit the GEV to the corresponding flood volumes (see Figure 2). It can be seen that there are also positive trends of the location and the scale parameter. The linear trend of the location parameter results in $s_{\mu,V} = 2.55 \pm 0.19 \cdot 10^7$ m³/a, whereas the linear trend of the scale parameter amounts to $s_{\sigma,V} = 1.06 \pm 0.08 \cdot 10^7$ m³/a. Both trends are significant on the 95% confidence level. The trend of the shape parameter results in $s_{k,V} = -2.29 \pm 1.13 \cdot 10^{-4}$ /a without a statistical significance.

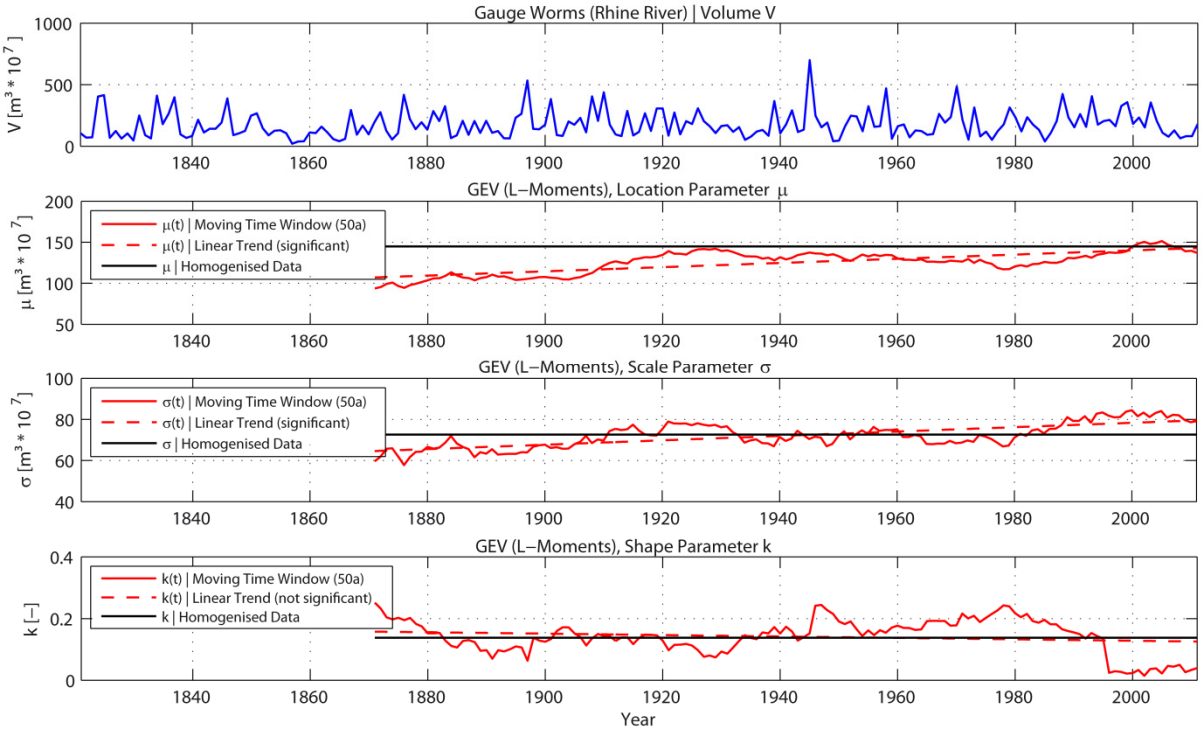


Figure 2: Same as Fig. 1 but for flood volume v

In a next step the nonstationarity of the dependence structure between the variables Q and V is investigated. Common measures for dependency are the correlation coefficients; here, Kendall's τ (Kendall, 1975), based on the ranks of two random variables is a widely used statistic to describe non-linear dependence.
Figure 3 illustrates the time dependent development of Kendall's τ between the AMAX discharges Q and the corresponding volumes V.

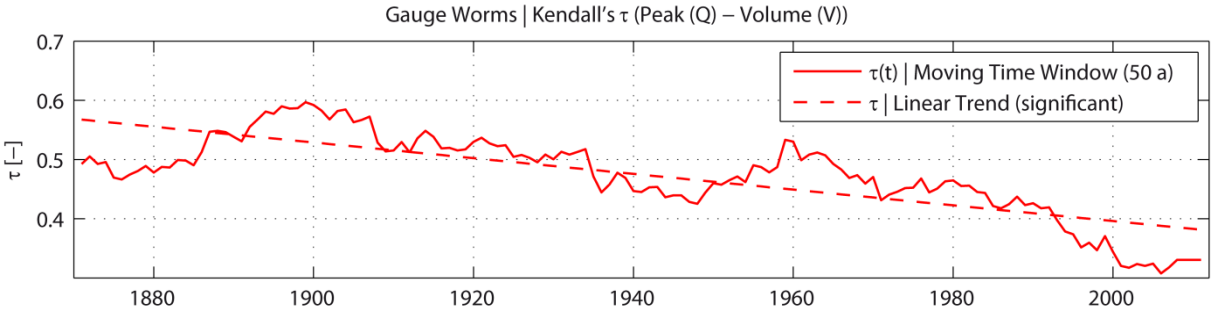


Figure 3: Time dependent Kendall's τ of annual maximum flood discharge (Q) and corresponding flood volume V derived with a 50-year moving time window (red solid line) and the significant linear trend (red dashed line)

Kendall's τ varies between a maximum value 0.59 at the end of the 19th century and a minimum of 0.31 at the beginning of the 21st century. We find a significant negative trend of s_τ = -0.0013 /a with a standard error of ±8.84 · 10^{-5} /a.
After proving the existence of nonstationarity in the marginals and in the dependence structure, we fit the copula functions outlined in Table 1 to the samples of Q and V resulting from calculating the moving averages. It is noted that the Ali-Mikhail-Haq copula is only defined for a range of Kendall's τ of [0,1/3). Hence, it is only considered where the time dependent τ does not violate this criterion, i.e. since 2001. For each copula we applied the parametric bootstrap procedure proposed by Genest et al. (2009). Although the Frank copula does not always yield the lowest S_n values, Figure 4 highlights that it leads to the best fit in the vast majority of cases.

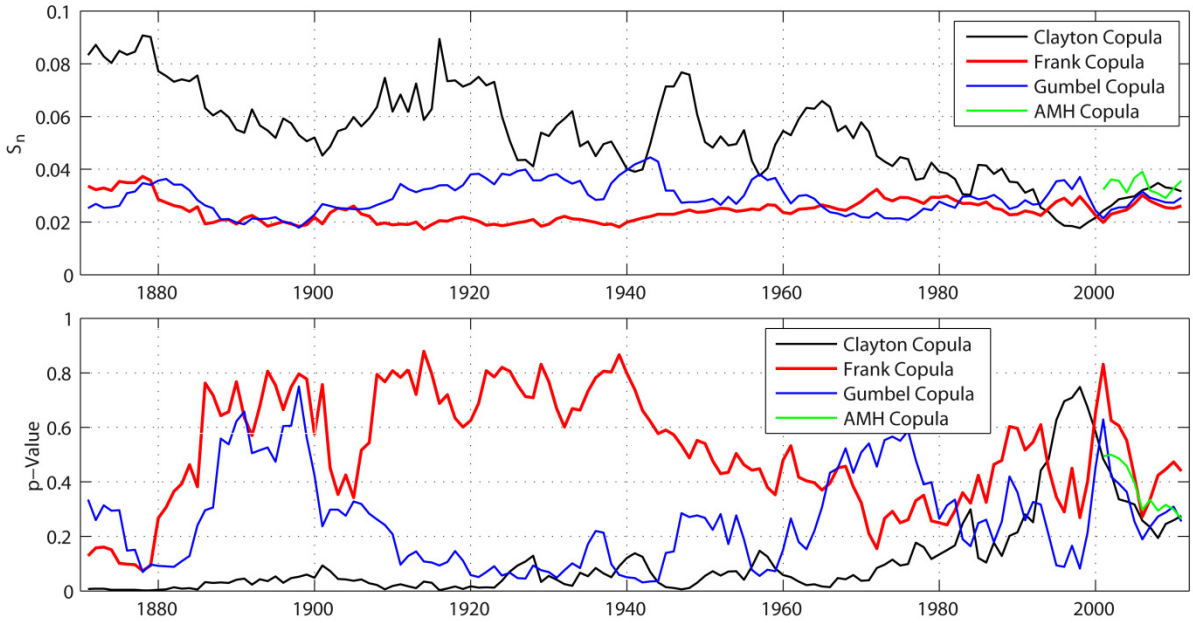


Figure 4: Results of the parametric bootstrap goodness-of-fit test. The upper figure illustrates the results of the Cramér-von-Mises statistic for the different Archimedean copulas. The lower figure shows the corresponding p-values

Furthermore the Frank copula is the only one of the four studied copulas with p-values above the 0.05 significance level for all time steps. Therefore, we choose the Frank copula and apply it to all 50-year time windows. It should be investigated in future works how a change of the marginal distributions and/or of the parametric copula can affect the result of the nonstationary bivariate statistical analysis.

4 RESULTS

With the time dependent parameters of the marginal distributions and the time dependent copula parameter θ(t) the temporal variation of the bivariate probabilities can be assessed for any desired event. In this study we focus exemplarily on events with a joint exceedance probability of $p_e = 0.02$. Although, in general, events with lower exceedance probabilities are of interest for hydrological practice, we chose this probability level to minimize the uncertainties of extrapolation. The results are illustrated by calculating the respective quantile-isoline.

Figure 5 shows the time dependent quantile-isolines for $P_E = 0.02$, beginning from 1871 (i.e. the 50-year time window from 1821 – 1871), for every time step up to 2011. The results in the upper left figure were derived from the nonstationary approach assuming all parameters of the marginals, i.e. μ, σ and k as well as the copula parameter θ as time dependent. The contour lines cover a broad range, with marginal values ranging from 4170 to 5490 m³/s for the discharge and 388 to 599· 10 m³ for the volume. Due to sudden changes of the magnitudes of the shape parameters of both marginals (e.g. Q in 1935 and V in 1994, ref. Figures 1 and 2) the iso-lines also cross each other.

In the upper right figure the location parameter μ and the scale parameter σ of the marginals as well as the copula parameter θ are assumed time dependent. The quantile-isolines cover a similar range as with nonstationary approach where all parameters are assumed time dependent. This is mainly caused by the strong influence of the scale parameters of both marginal distributions. However, there are obviously less contour lines crossing. Generally, there is a clear increase of the univariate values. While in 1871, for example, the $P_E = 0.02$ peak discharge amounts to 4299 m³/s, in 2011 it amounts to 5761 m³/s.

For the sake of comparison the quantile-isoline derived from the homogenized (i.e. de-trended and assumed stationary) data is also illustrated.

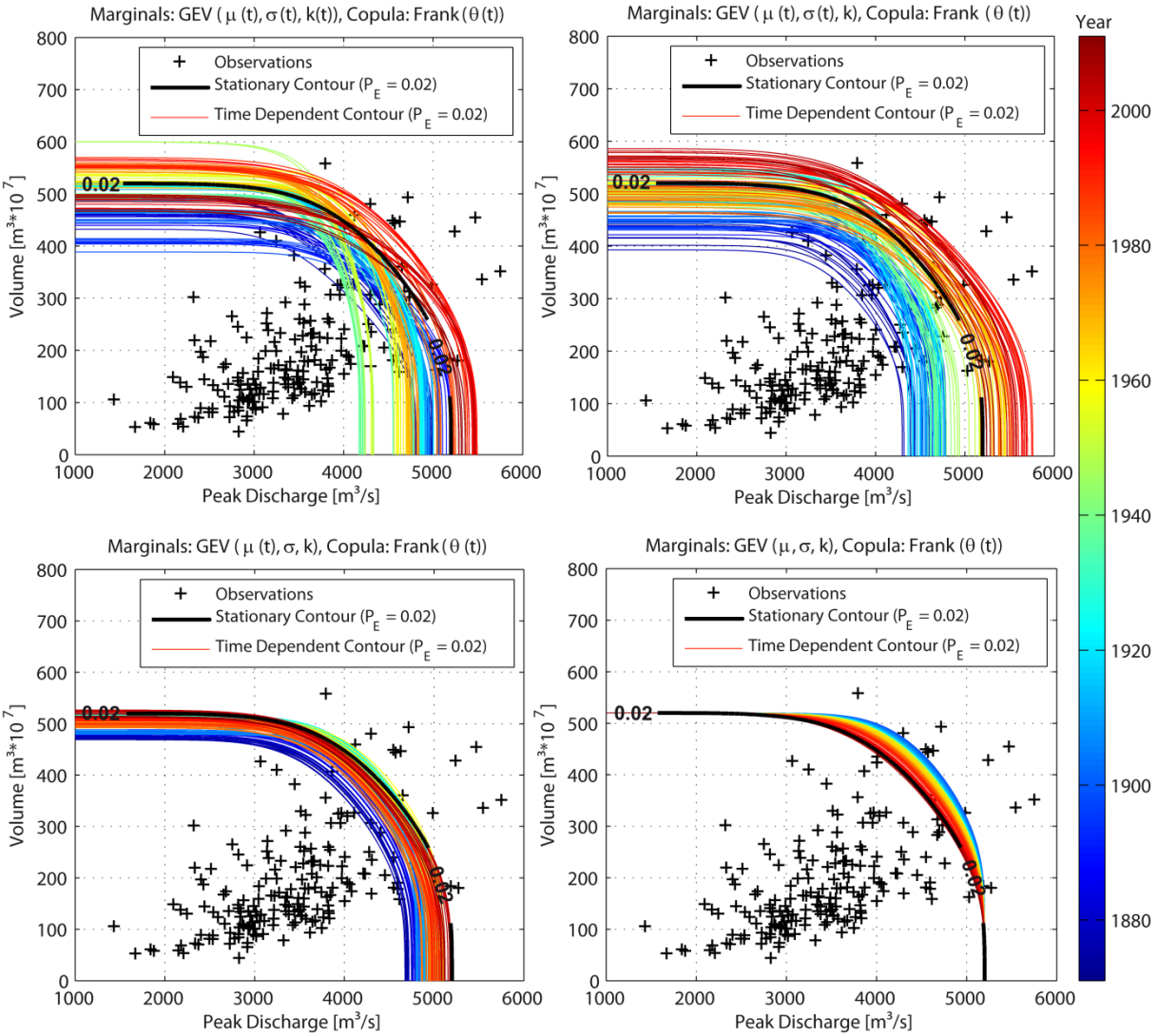


Figure 5: Time dependent quantile-isolines (P_E = 0.02) for the cases μ(t), σ(t), k(t), θ(t) (upper left), μ(t), σ(t), k, θ(t) (upper right), μ(t), σ, k, θ(t) (lower left), μ, σ, k, θ(t) (lower right)

In the lower left panel only the location parameter μ and the copula parameter θ are time dependent. The contours cover a much smaller range,, but a positive shifting is still evident.
To illustrate the influence only of the temporal variation of dependence between Q and V, Figure 5 shows in the lower right panel the development of the quantile-isoline with stationary marginals and a time dependent copula parameter θ. In this case the quantile-isolines move to the lower left over time, as a result of the decreasing correlation coefficient that has already been discussed above.

5 DISCUSSION AND CONCLUSION

In this paper, a bivariate nonstationary approach is introduced to investigate the time dependent behavior of bivariate design parameters. The approach is based on modeling the marginals using the GEV distribution and Archimedean copulas for model the dependence structure. A 192 years long discharge time series was considered to show, exemplarily, that bivariate design parameters (here Q and V) can change significantly over time, due to trends in the marginal distribution parameters and changes in the dependence structure between the variables under investigation. In future studies those trends should be extrapolated in order to assess design events for future time horizons.
It is noted that in this present study neither the type of the marginal distributions nor the copula type changed over the time. It is very likely that the application of this approach to other data sets will show different results.

6 REFERENCES

bfg (2012): *Online gauge information system*, http://undine.bafg.de, last visit on website: 12 September 2013.

Chebana, F., Ouarda, T.B.M.J. and Duong, T.C. (2013): *Testing for multivariate trends in hydrologic frequency analysis,* Journal of Hydrology 486, pp. 519-530.

Coles, S. (2001): *An Introduction to Statistical Modeling of Extreme Values*, Springer, ISBN 1-85233-459-2.

De Michele, C. and Salvadori, G. (2003): *A generalized Pareto intensity duration model of storm rain-fall exploiting 2-copulas*, J. Geophys. Res. 108 (D2), pp. 1–11.

Favre, A-C., Adlouni, S., Perreault, L., Thiémonge, N. and Bobée, B. (2004): *Multivariate hydrological frequency analysis using copulas*, Water Resour. Res. 40 (1), W01101.

Genest, C., Rémillard, B. and Beaudoin, D. (2009): *Goodness-of-fit tests for copulas: A review and a power study*. Insurance: Mathematics and Economic, 44, pp. 199-213.

Grimaldi, S. and Serinaldi, F. (2006): *Design hyetographs analysis with 3-copula function,* Hydrol. Sci. J. 51 (2), pp. 223–238.

Hosking, J.R.M. and Wallis, J.R. (1997): *Regional frequency analysis.* Cambridge, UK: Cambridge University Press.

Kao, S. and Govindaraju, R. S. (2010): *A copula-based joint deficit index for droughts*, J. Hydrol. (Amsterdam) 380 (1–2), pp. 121–134.

Karmakar, S. and Simonovic, S. P. (2009): *Bivariate flood frequency analysis. Part 2: A copula-based approach with mixed marginal distributions*, J. Flood Risk Manage. 2 (1), pp. 32–44.

Kendall, M.G. (1975): *Rank Correlation Methods*, Griffin, London.

Mann, H.B. (1945): *Nonparametric tests against trend*, Econometrica 13, pp. 245 – 259.

Mudersbach, C. and Jensen, J. (2010): *Nonstationary extreme value analysis of annual maximum water levels for designing coastal structures on the German North Sea coastline*, J Flood Risk Management 3, pp. 52-62.

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

Schölzel, C. and Friederichs, P. (2008): *Multivariate non-normally distributed random variables in climate research – introduction to the copula approach*, Nonlinear Proc. Geophys. 15, pp. 761–772.

Shiau, J. T. (2006): *Fitting drought duration and severity with twodimensional copulas*, Water Resour. Manage., 20 (5), pp. 795–815.

Sklar, A. (1959): *Fonction de répartition à n dimensions et leurs marges*. Publications de Institut de Statistique Université de Paris, 8, pp. 229–231.

Song, S. and Singh, V. P. (2010a): *Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data*, Stoch. Environ. Res. Risk Assess. 24 (3), pp. 425–444.

Song, S. and Singh, V. P. (2010b*): Frequency analysis of droughts using the plackett copula and parameter estimation by genetic algorithm*, Stoch. Environ. Res. Risk Assess. 24 (5), pp. 783–805.

Wahl, T., Mudersbach, C. and 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, pp. 495-510.

Zhang, L. and Singh, V. P. (2006): *Bivariate flood frequency analysis using the copula method*, J. Hydrol. Eng. 11 (2), pp. 150–164.

Zhang, L. and Singh, V. P. (2007): *Gumbel Hougaard copula for trivariate rainfall frequency analysis*, J. Hydrol. Eng. 12(4), pp. 409–419.

Inundation caused by dike break – real-time forecast and monitoring during the Flood 2013

Robert Juepner[1] and Thilo Weichel[2]

[1]University of Kaiserslautern, Institute for Hydraulic Engineering and Water Management, Kaiserslautern, Germany, Email: robert.juepner@bauing.uni-kl.de

[2]State authority for flood protection and water management Saxony-Anhalt, Halle, Germany

Abstract

The failure of a dike during a catastrophic flood event usually leads to inundation of large areas. Such an event can be analysed using two-dimensional (2D) hydraulic models to describe the size of the inundation area, water depths as well as flow velocities. The model usually requires a precise Digital Elevation Model (DEM) together with detailed information about the dike breach itself, specifically, the geometry of the breach and discharge.

During a catastrophic flood event, the required hydraulic and hydrologic information necessary to describe the physical extent of the dike breach is usually not available or missing. Furthermore the emergency management agencies usually need reliable information about the dike breach and inundation in a very short time frame in order to take the appropriate actions, such as implementing appropriate evacuation measures. Such requirements from emergency management agencies are often a big challenge for water management experts taking responsibility during catastrophic flood events.

For example, during the recent Elbe flood disaster in June 2013, a catastrophic dike breach occurred near the village of Fischbeck in Saxony Anhalt, Germany. As a result, nearly 500 m³/s of flood water was flowing into relatively flat areas behind the dike through a 90 m breach. The responsible emergency management forces tried to record document and forecast the development of the flooding with help of suitable methods. Because of the complex set-up and modelling requirements of hydraulic models especially in floodplains, the application and use of geographic information systems (GIS) was realised. As a first step, an area-wide DEM with a high resolution was spatially aggregated to a practical 2D-grid level of 5 m x 5 m. Based on this data set and water levels (which were directly measured in the flooded zone) different maps of the actual and potential flood extent were computed using readily available GIS computational tools. Incorporating the known slope of water surface in the main stream, usable results of real time forecast and monitoring of flood levels could be made.

Additionally, a daily surveying flight was undertaken to monitor the inundation area and to receive real-time information about the flow velocity as well as flow direction. As a result, a daily map illustrating the inundation area as well as the predicted inundation area was created. Together with the results of the GIS-computing, the flood recording and monitoring via aerial documentation supported the management expert in case of such an extreme situation. The maps were used as a basis for emergency management actions and the final situation of the dike break flooding could be nearly forecasted.

1 INTRODUCTION

In June of 2013, a devastating flood event occurred in the Elbe catchment as well as the upper Danube in Bavaria. It resulted from heavy rainfall events from May 17th until June 2nd combined with extreme values of soil moistures. Spanning more than 250 km along the Elbe river, the highest recorded water levels were exceeded (BfG, 2013).

Figure 1: Dike at the river Elbe close to Fischbeck on June, 9 – 10 hours before the break (Ellmann, 2013)

On June 10th 2013 at nearly midnight, a 50 m wide breach in the dike developed along Elbe River close to the village of Fischbeck north of Magdeburg expanding to a final width of 90 m. At that time, the water level at the Elbe was measured to 8.36 m at Tangermuende gauging station. This level was 68 cm above the highest recoded water level in the river with nearly no remaining free board at the dike.

The reason for the dike break is still subject for investigation. It was obvious, that both inner and outer slopes of the dike became unstable (see Figure 1). As a result, the dike crest was sinking a few decimeters. Even though the emergency management forces as well as numerous volunteers tried to heighten the crest with sand bags for several hours, they were still unsuccessful in the end. The dike crest was overtopped in this area and the entire dike was eroding very fast (Figure 2).

The discharge through the breach was estimated approximately at 500 m³/s in the beginning. After a period of four days another discharge measurement was taken at 320 m³/s. Due to the catastrophic results of the dike break, a drastic approach was undertaken to close the breach with the help of ships. Three prams were countersunk within a six day period after the breach started. This allowed the construction of an artificial contrivance dike to stop the flooding of the back-country as illustrated in figure 3. It was estimated, that during the dike breach, an overall amount of nearly 225 Million m³ of water from the Elbe River flooded the area behind the dike.

Figure 2: Dike break at the river Elbe close to Fischbeck on June, 10 – 18 hours after the event (Juepner, 2013)

Figure 3: Dike break at the river Elbe close to Fischbeck on June, 22 – 12 days after the event (Mueller, 2013)

2 EFFECTS OF THE DIKE BREAK

The area protected by the dike was flood prone and comparably flat. From a geomorphological point of view this area is part of the initial stream channel of the Elbe River. The flood water was moving parallel to the Elbe river for roughly 30 km following the slope of the valley terminating in river Havel (Figure 4 and 5).

The huge amount of water – more than 40 million cubic meters within the first day – were inundating a large part of the flat land as shown in Figure 4 below. Even if this area was not heavily populated, a couple thousands of people's lives were endangered and needed to be evacuated. Nearly 150 km² of land had been inundated and caused extensive damage for the effected people.

For the responsible emergency management it was very difficult to forecast the main direction of the water flowing as well as the temporal development of the inundation. Therefore this information was very much needed to develop some sort of evacuation strategy and for being able to mitigate the negative effects of the inundation.

Figure 4: Part of the inundated area beside the river Elbe on June, 11 – 36 hours after the event (Juepner, 2013)

3 REAL-TIME FORECAST AND MONITORING FOR THE EMERGENCY MANAGEMENT

The 2013 flood along the Elbe downstream of the confluence with river Saale was characterised by water levels never observed in the past. Table 1 shows the difference between the maximum water level during the flood in June of 2013 and the past historical peaks at selected gauges. Differences of more than 60 cm were the reason of this catastrophic flood event and the dike breaks in the affected area. In such an extreme situation the coordination of the flood emergency management to implement real time hazard control measures is very difficult, especially in this particular case with new water level records. There are nearly no existing "flood scenarios" available to deal with this complex flooding situation.

Table 1: Maximum peak of water level during the flood in 06/2013

Gauging station	maximum peak [m]	historical peak [m]	difference [m]
Tangermuende	8.36	7.68	+ 0.68
Magdeburg - Strombruecke	7.46	6.80	+ 0.66
Barby	7.61	7.33	+ 0.28

Table 2 compares the observed maximum water levels in the recent flood of 2013 along the Elbe river with the estimated (based on past historical flood levels) 100 and 200 year return period floods and illustrates the uniqueness of this flood event. Unfortunately the existing technical flood protection measures are designed to the 100 and 200 year return period flood levels.

Table 2: Maximum peak 06/2013 and water levels of return periods

Gauging station	maximum peak [m]	HQ_{100} level [m]	HQ_{200} level [m]
Tangermuende	8.36	8.00	8.28
Magdeburg - Strombruecke	7.46	6.95	7.13
Barby	7.61	7.06	7.24

In case of a dike breach, the main information required by the emergency management forces are:

- spatial and temporal enlargement and progress of the flooding,
- direction and intensity of the flood wave and the size of the expected inundation area,
- Differences between forecasted water levels and the elevations of flood protection measures (dikes, walls sand bag barriers, etc.).

To provide such information, it's necessary to create appropriate maps about the spatial extent of the flood, water depths, flow direction, freeboard of dikes (see Figure 5) or evacuation routes and other appropriate measures. The experiences from this flood event showed once again that maps together with a plot of digital data (Juepner, 2003) are the best kind of documents for any emergency management agency during a catastrophic flood event. Beside the use of maps in the real-time forecast of potential flood situations, the recording and monitoring of a flood event is to be made by maps.

Figure 5: Map with classification of freeboard along Elbe dike by scenario with a return period of 200 years (LHW)

Until the extreme flood in June 2013, an existing hydraulic 2D-model for the river Elbe pointed out that discharges with a probability up to the valid return period of 200 years would not cause any damage to the dike even if the freeboard would very small during the flood. Because of this fact, the spatial extent of the existing model was limited by the line of the dikes (see Figure 5). Therefore no hydraulic model-mesh was available outside the line of the dikes as well as the flooded area behind the dike break near Fischbeck (Figure 6). It is generally difficult in practice to generate detailed 2D-model-meshs for dike break scenarios with a potential flooded area of more than 100 km².

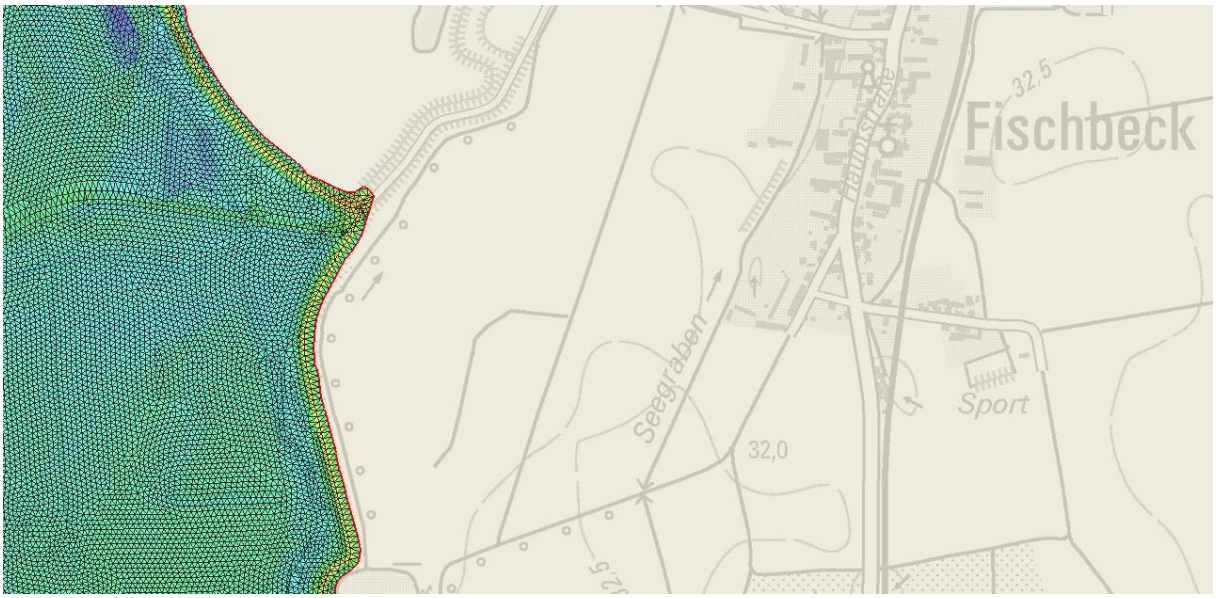


Figure 6: Extend of the existing 2D-model mesh limited by dike line near Fischbeck (LHW)

Consequently, this flooding case required alternative methods and tools, like the hydraulic model presently in use referred to as FLOODAREA (Assmann, 2005). This kind of modelling approach, which is based on coupled GIS and hydraulic components, was used in the preparation of the Elbe-Atlas for assessing the inundation area associated with dike failure (Assmann et al., 2006). The approach focused on the input of DEM-data. Further investigation of dike break scenarios in this region was made with this tool by Ettmer (2011). The consequences of dike break scenarios along the left embankment of the Elbe river were investigated together with the impact of flood enlargement to the city of Stendal. But without any data preparation the set-up of the model parameters such as the drainage system and soil moisture accounting as required to run the model, is not a realistic option. Furthermore, the computational time required by the model makes it difficult to implement in real time during the flood.

Therefore alternatives for a real-time forecast of inundations caused by dike break were required during the Elbe flood disaster in June 2013. No existing flood scenarios were available for the responsible water management authority. In this context the usage of DEM-data proved to be extremely valuable and important basic data for further investigations managing the flood in real time. The availability of an area-wide detailed digital elevation model (DEM) with a spatial resolution of 1m x 1m and vertical precision of 0.15 m was used to derive information about the topography and the potential flooding situation. As a first step the original DEM (1 m x 1 m) was aggregated to a DEM5 with a spatial resolution of 5 m x 5 m to permit a better handling and performance. The deficit to a DEM with a resolution of 1 m x 1 m was neglected in this kind of usage. Finally, the application of Geographic Information Systems (GIS) together with this topographic data, it was possible t to visualise and compute information needed by the flood emergency management authorities (refer to Figure 7 below).

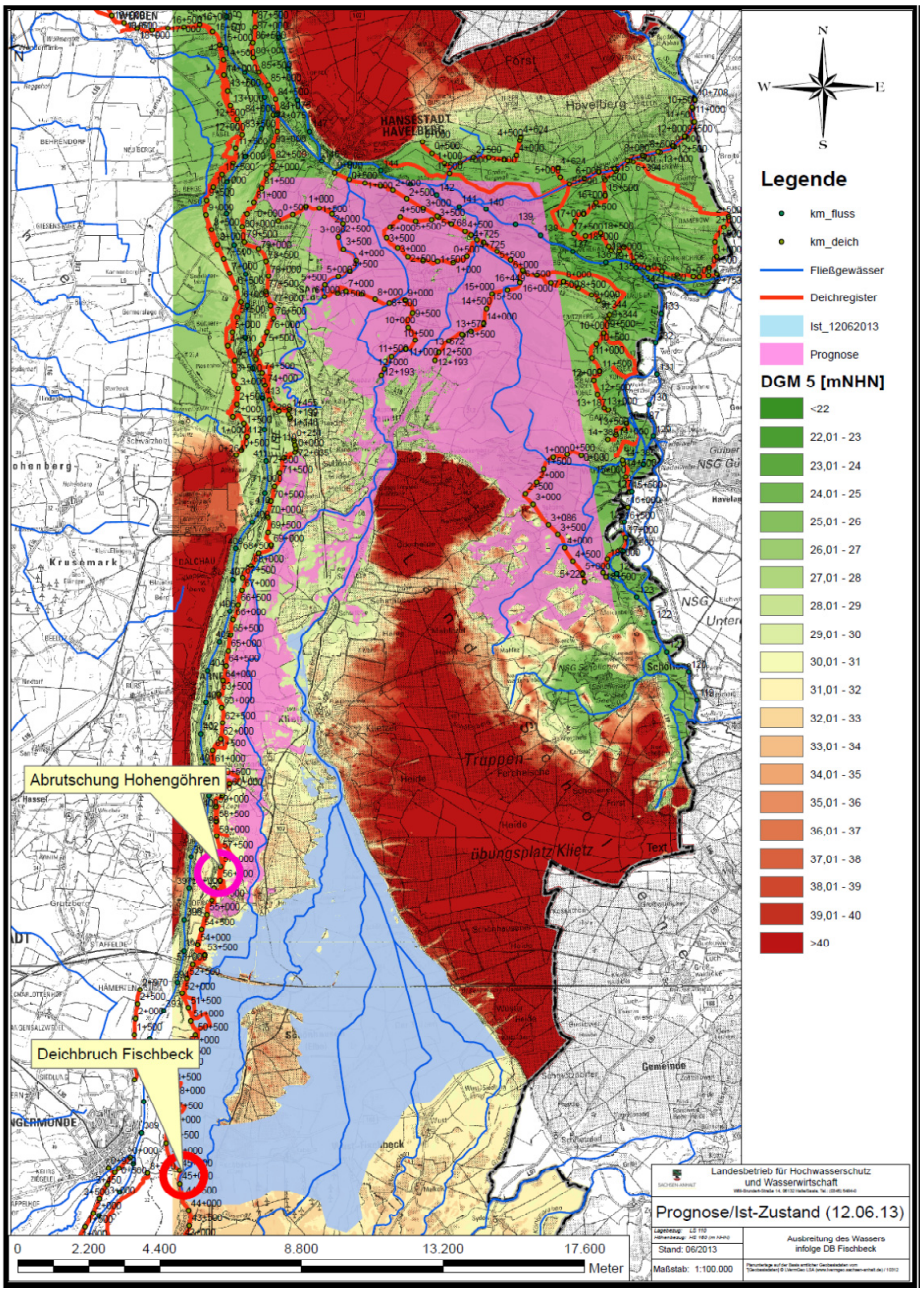


Figure 7: Map illustrating the existing inundated area (light blue) and forecasted inundated area (pink) on June, 12 (Juepner, Mueller and Weichel, 2013)

Local water levels in the flooded area were measured at different times after the dike break along the water-land-line. Considering the known slope (Figure 8/1) of the water surface (adopted from the already existing hydraulic modelling of the Elbe), the surveyed water levels were recalculated in an area wide GIS-dataset (Figure 8/2). An additional new GIS raster dataset, which represented the theoretical water surface, was created with the same extent and spatial resolution of the DEM5. Because of the same spatial extent, the intersection of both raster datasets could be realised (Figure 8/3). In the next step the difference between DEM5 and the GIS-dataset of the theoretical water surface, based on local measures, was calculated (Figure 8/4). The result of the intersection was a new GIS-dataset with potential inundation area and water depths based on a 5 m x 5 m resolution.

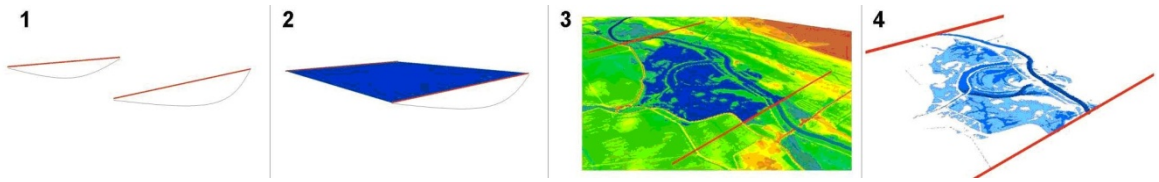


Figure 8: GIS-computing of potential flooding areas (Weichel, 2011)

The result of the intersection described above was drawn on a map and covered an area of more than 30km from the origin of the dike break (near Fischbeck) to the city of Havelberg (Figure 7). Based on this information the large-scale situation of the dike break flooding could be illustrated. The main advantage of this simple GIS-intersection over other methods is the relative short time to compute the results. In such extreme situation, like the 2013 flood, no detailed information about culverts, bridges or the stream network are necessary, if the primary interest is on the forecast of inundated areas caused by dike break.

Because of some missing detailed information in the results presented above and to evaluate their accuracy, further investigation was carried out using daily observation flights over the flooded area. By using helicopter an airborne based data acquisition was realized with this aerial reconnaissance. Beside aerial pictures, the delineation of the real flooded area was documented daily as well as being video recorded by the on board technology of the German Federal Police (Bundespolizei). The additional information of the aerial reconnaissance was used to create and update daily inundation maps as an instrument for the real-time recording and prediction for the emergency management.

Figure 7 illustrates a map of the actual flooded area on the 12th of June and the prediction of the further flooding, created by the GIS intersection. Figure 9 shows the detailed situation of the flooded areas around the village of Hohengoehren using water depths calculated using GIS modeling and actual measurements.

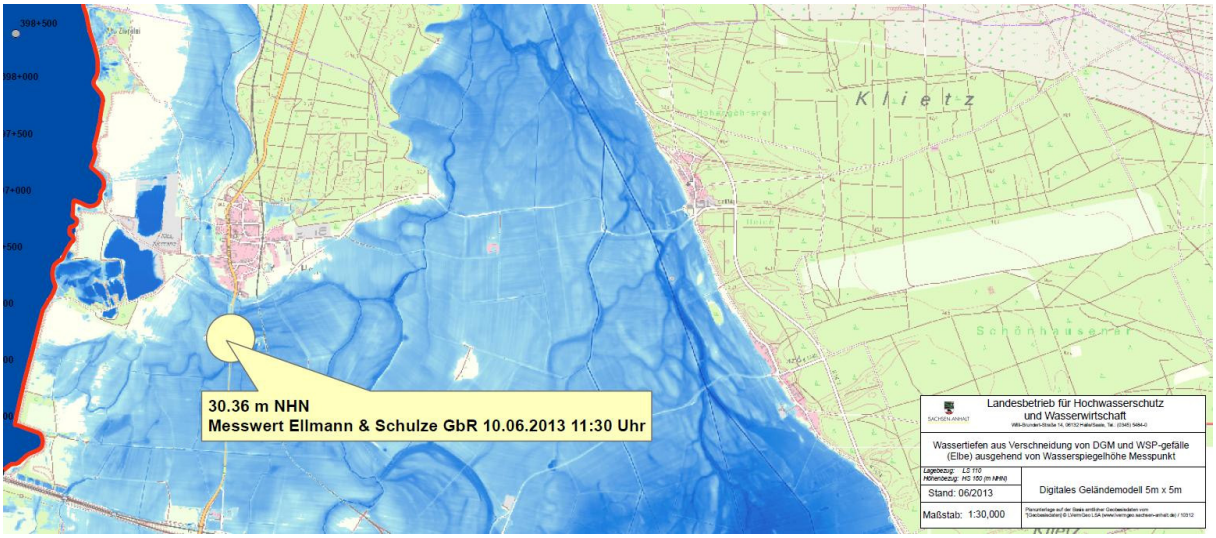


Figure 9: Map of water depths calculated by the intersection of DEM5 and water surface based on local measurement

In summary, the following results were compiled in this study:

- Flood inundation maps were created daily during the emergency management of the flood of June 2013 along the Elbe river for the inundated area as a result of the dike break near Fischbeck. They were created based on real-time data-acquisition and combined modeling by GIS-intersection.
- A comparison was possible of computed flood inundation with observed satellite images showing the inundated area with a delay of more than 24 hours
- Beside this, a daily report on the flooding situation was created and illustrated with aerial pictures and was used by emergency management forces.

4 DISCUSSION

In an extreme situation like a dike break during a flood, relevant information is extremely valuable for the affected people as well as for decision-makers in the emergency management positions. Such information is essential yet characterized by uncertainties. The main reasons for the uncertainties are based on the spatial and temporal description of the flood affected area. The dike break near Fischbeck and with the subsequent flooding of an area of nearly 150 km² was characterized by uncertainty and a new situation for all the stake holders of this natural disaster. Therefore it was very important to provide, as a part of the emergency management process, reliable information in order for the emergency management personnel to take the necessary actions.

The use of the GIS-based intersection of DEM data and local water level measures has been demonstrated as a successful approach for delineating the flood inundation areas in this study. The results have been of good quality. Different situations of flooding could be well predicted, because the precision of the high resolution DEM provide area wide information about heights. Even if the resolution in an aggregated DEM5 is reduced compared with the original DEM1, it's possible to derive useful information about flow direction and water levels. Also the transfer of a known water level decline in such a steady flow situation can be regarded as a pragmatic approach. So the situation near the village Kamern (Figure 9), 30 km away from the point of break, could be well forecasted. Figure 10 shows the prediction with water flow in different lines, overflowed roads and the inflow behind the dike line. During the aerial observation of the affected area the details of flooding as shown in figure 10 was confirmed.

Figure 10: Predicted inundation areas near Kamern

Real problems with this kind of prediction resulted in the installation of local flood protection measures, like barriers of sand filled big bags with heights of approx. 1 meter. Because of these barriers, the hydraulic situation was modified and cannot be predicted with this GIS-projection.

Compared with the possibility of a detailed hydraulic 2D-modelling, the described methodical approach used the same topographic data (DEM) and the upper boundary condition (local water level measure). In addition to that, the information of the water level decline, which based on results of a steady flow modeling of the adjacent Elbe as well as daily aerial observations have been used to verify the predictions. Problems occurred by the usage of the aerial photos were mostly related with areas of an intensive vegetation (forest, fields, etc.). In this region it was very difficult to really observe the delineation of the water-land boundary.

No practical examples could be used as a "roadmap for real-time forecast for inundation areas". But the combination of the already known facilities of GIS-tools, the existing detailed data and the

cooperation of different stake holders affected by the flooding event led to results that adequately met the requirements of a flood emergency management initiative.

5 CONCLUSION

In June 2013 the dike break at the Elbe River near Fischbeck threatened a large area with flood inundation. The requirement of the emergency management forces was to get reliable data as soon as possible on the effects flooding as a result due to the dike break. This was a necessity for the protection of the people who were suffering a lot by this situation.
The water management and flood protection authority was not provided with flood scenarios describing this situation neither appropriate model calculations had been made before the event itself. Therefore it was necessary to find and adapt a methodology to deliver reliable information to cope with this catastrophe. The described combined approach of using a GIS-based intersection of DEM data and local water level measures and receive data by a daily airborne surveying by helicopter was able to get appropriate information for the emergency management forces. Even if this approach is subject to further investigation there is an obvious gap in scientific research on “real-time monitoring” of catastrophic events like dike breaks at large rivers.

6 ACKNOWLEDGMENTS

The paper describes first results from emergency management during the Elbe flood disaster in June 2013 at the Elbe River. Most data have been taken from the emergency management of Saxony-Anhalt (f.e. Hochwasser-Einsatzstab des LHW). We would like to mention that many engaged experts worked quite hard to cope with this catastrophe during that time and we do have the privilege to publish some of the first results to the scientific community. In addition to that we would especially like to thank Stefan Mueller from the University of Applied Sciences Magdeburg who was our partner during the entire crisis. Last but not least: without the fine work of the helicopter pilots of the German Federal Police and their very cooperative support we wouldn’t have been able to get the information we needed for our work.

7 REFERENCES

Assmann, A. (2005): *FloodArea, ArcView-Erweiterung zur Berechnung von Überschwemmungs-bereichen*. Software-Handbuch. Heidelberg [in German].

Assmann, A.; Grafe, M.; Runge, I. and F. Thäger (2006): *Einflüsse des Berechnungsverfahrens und der Qualität der Grundlagendaten auf die Ermittlung überschwemmungsgefährdeter Gebiete.* Hydrologie und Wasserbewirtschaftung, H.1, 19-24 [in German].

BfG (2013): Bundesanstalt für Gewässerkunde (BfG): *Das Juni-Hochwasser des Jahres 2013 in Deutschland*, BfG-1793, Koblenz [in German].

Ettmer, B. (2011): *Jahresbericht 2011 des Instituts für Wasserwirtschaft und Ökotechnologie* (IWO. Magdeburg, [in German].

Jüpner, R. (2003): *Erfahrungen aus dem operativen Hochwasserschutz:* WASSER und ABFALL, 5/2003 [in German].

Weichel, T. (2011): *Entwicklung eines Werkzeugs zur systematischen Bewertung der Grundlagen von Hochwassergefahrenkarten*. Dissertation. Fachbereich Bauingenieurwesen der Technischen Universität, Kaiserslautern [in German].

Extreme flood events in the Elbe river catchment – How reliable is the 'traditional' parameter HQ_{100} for designing flood protection measures?

Christoph Mudersbach[1]

[1]Research Institute for Water and Environment, University of Siegen, Siegen, Germany, Email: christoph.mudersbach@uni-siegen.de

Abstract

In this study we investigated, how the 100-year level at the gauges Neu Darchau and Dresden is affected by the latest extreme events in the years 2002, 2006, 2011, and 2013. It was shown, that the 100-year flood event at the gauge Neu Darchau is less affected by the extreme events than at the gauge Dresden. Additionally, from record theory the result was obtained, that there is a 83% chance, that the 2013 record is rather due to non-stationary changes in the system than due to the natural variability.

1 INTRODUCTION

The design of flood protection measures is dominated by using the 100-year flood event (HQ_{100}). The 100-year flood event is a statistically derived discharge, which will be reached or exceeded once in 100 years on average. This value is not only used in Germany, but in many countries worldwide, e.g. the United States, where it is known as the 1 % chance flood standard. The 100-year flood event has a long history, dating back at least until 1968. In this year, the 100-year flood event was established in the United States with the National Flood Insurance Act (FEMA, 2013). This level '*was selected because it was already being used by some agencies, and because it was thought of that magnitude and frequency represented both a reasonable probability of occurrence and loss worth protecting against and an intermediate level that would alert planners and property owners to the effects of even greater floods'* (National Flood Policy Forum, 2004). In later years, reforms of the National Flood Insurance Act in 1976 and 1983 stated, that there is no justification of converting to another standard.

Within the last few decades, as experience in using the 100-year flood level increased, some concerns in the United States have arisen: (1) Floods appear to be bigger and causing more damages than expected, (2) advanced geospatial technology, and modeling technology that were not available several years ago call into question the reliability of the flood standard, (3) the uncertainties in estimating the 100-year flood levels are rather high. In 2012, the Biggert-Waters Flood Insurance Reform Act (FEMA, 2013) extended the flood mapping program to map all areas within the 100-year and 500-year floodplains.

In Germany the 100-year flood event also serves as a basis for the design of flood protection measures since several decades. The latest revision of the German guideline for flood protection works on rivers (DIN 19712:2013-01) recommend a 100-year flood level for the protection of human settlements and industrial facilities. Higher return levels are only recommended for objects, where extraordinary consequences in case of a flood event could be expected. The respective return level has be to defined individually in such cases, while return levels up to 500-years are used in practice. In this context it is interesting to mention, that the European flood directive (2007/60/EC), which is also adopted into national law, defines a 100-year flood as a medium probability event and so-called extreme events are rated in the class with low probabilities. This simply means, a 100-year event is not classified as an extreme event. In Germany there is no unique definition on how to determine extreme events in accordance to the European flood directive. In some federal states it is defined as a 200- or 500-year event, whereas in other federal states an extreme event is defined as 1.5 to 2.0 times the 100-year flood event.

In addition, the public view on extreme events is also a little bit different. If an extreme flood occurs, this is quickly named as a one-in-a-century flood. Such events are sometimes also treated as equal with record floods or highest possible floods. That means the public perception and understanding of these terms are far away from 'medium probability events' as defined in the European flood directive.

During the last decades several severe floods occurred in different river basins in Germany (e.g. 1993, 1995 Rhine; 1997 Odra; 1999, 2001, 2006, 2013 Danube; 2002, 2006, 2011, 2013 Elbe). In a public perception there seems to be a trend to more frequent hydrological extreme events resulting in a rising flood hazard. Thus, the question arise how much the traditional design parameter HQ_{100} is affected by these events and if it can be assumed as a reliable basis for future design purposes?

2 DATA

Within this investigation, we focus on analyses of the discharge data of the gauge Neu Darchau and Dresden (Elbe River). The gauge Neu Darchau is the most downstream discharge gauge of the Elbe river (Elbe-km 536.4) before it becomes an estuary. The gauge is considered as a benchmark for the total discharge of the Elbe river. The gauge Dresden thus represents the upper Elbe (Elbe-km 55.63) and is one of the best known gauges at the Elbe river.

The Neu Darchau data were obtained from the Water and Shipping Office Lauenburg (WSA Lauenburg), which is the official gauge operator. The records comprise daily mean discharge values from 1 November 1874 to 31 July 2013 (hydrological year in Germany: 1 November to 31 October), resulting in a time series covering 140 years without any gaps. The discharge measurements operate regularly since 1874 without any discontinuities. Since 1 November 1997 data with a resolution in time of 15 minutes are available. Prior to that, several measurements per day (not equally distributed) are the basis for the daily mean discharge data. Since the catchment size amounts to 131,950 km², these daily measurements are also suitable to compute representative daily mean discharge data.

The most extreme flood in the official data set is on 25 March 1888 with HHQ = 4400 m³/s. Since it was a severe winter flood, it is known that during this event river icing occurred in the Elbe river which significantly influenced the flood stage measurement at gauge Neu Darchau (WSA Lauenburg, 2012). Hence, the dedicated discharge of 4400 m³/s is also affected by river icing. For this flood event the WSA Lauenburg also provides a corrected peak discharge of 2310 m³/s on 24 March 1888 instead of 4400 m³/s on 25 March 1888 (Rölver, 2012), which allows for the incorrect flood stage measurement. Although the corrected discharge is not officially fixed, we decided to use this corrected value instead of the original value, since the original –and obviously incorrect– value would significantly affect the extreme value statistics.

Figure 1 (top) illustrates the corrected daily mean discharge series at gauge Neu Darchau from 1875 to 2013 (July). The overall mean of the time series amounts to 711 m³/s with a standard deviation of 447 m³/s. Within the last years several extreme events occurred, with the following measured discharges: HQ_{2002} = 3410 m³/s, HQ_{2006} = 3590 m³/s, HQ_{2011} = 3593 m³/s, and HQ_{2013} = 4060 m³/s. All events belong to the 11 highest flood events measured. Moreover, the 2013-event was a new record.

The Dresden data were obtained from the Water and Shipping Administration (BFG) and comprise daily mean discharge data from 01 November 1806 to 31 July 2013, resulting in a time series covering 208 years without any gaps. The catchment size of the gauge Dresden amounts to 51,400 km². It has to be mentioned, that both the data for June and July 2013 from the gauges Neu Darchau and Dresden are preliminary and not officially validated.

The discharge time series of the gauge Dresden from 1807 to 2013 is shown in Figure 1 (bottom). The mean of the overall time series is 334 m³/s with a standard deviation of 283 m³/s. From the extreme events within the last years only two floods are ranked within the list of the 10 highest floods: HQ_{2002} = 4500 m³/s and HQ_{2013} = 4350 m³/s.

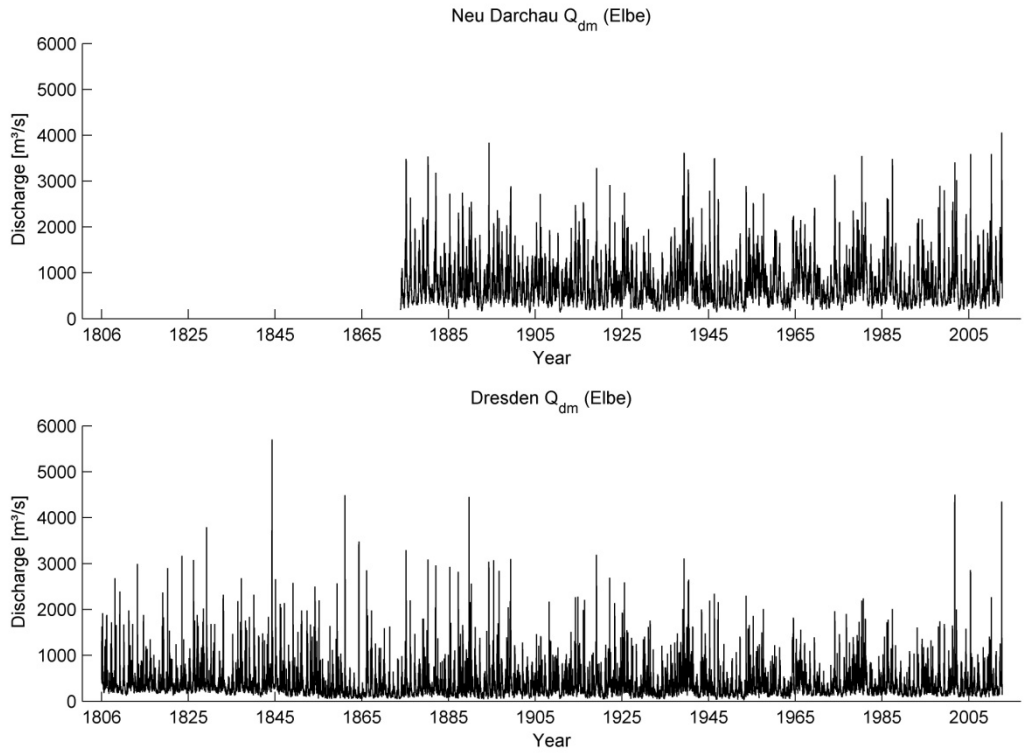


Figure 1: Time series of daily mean discharge values of gauges Neu Darchau (top) and Dresden (bottom)

3 METHODS

Six flood indicators were analysed in this study using both, the block maxima and the peak-over-threshold approach. The most common flood indicator in flood trend studies is the annual maximum discharge, i.e. the largest daily mean discharge that occurs in each hydrological year. This flood indicator is labelled as AMF. The annual maxima approach has extensively been used in the past (Acero *et al.*, 2011). However, it can be a wasteful method if further data of extremes are available (Coles, 2001). Conversely, if no extreme flood occurs within a year, the maximum value will still be selected. To overcome these shortcomings, some alternative approaches came up in hydrological statistics. The most prominent methods are the r-largest approach (e.g. Coles, 2001) and the peak-over-threshold (POT) approach (e.g. Leadbetter, 1991; Bayliss & Jones, 1993; Coles, 2001).

In the r-largest approach, not only the annual maximum values (r = 1) are considered in the sample, but e.g. the two (r = 2) or three (r = 3) largest annual values. The advantages and disadvantages of this method are obvious. Given a year with several extreme floods, using the r-largest method extends the data basis by including more of the available information concerning extreme discharge events. In contrast, if a year has no major floods, using the r-largest approach still considers the r-largest events of this year within the sample. In this study we compute the annual r-largest samples considering the r = 2 and r = 3 largest events per year, hereafter referred to as AMFr2 and AMFr3. Discharge datasets can exhibit dependencies that are related to the same event that caused these floods. By creating a sample of the r-largest values per year, one has to ensure independence of the selected events, which means that the events should have a certain distance in time (declustering time). Following Svensson *et al.* (2005), we use a declustering time of 20 days.

The POT approach (also known as partial duration series) provides a more flexible representation of floods compared to the AMF approach, since it accounts for stochastically and unequally distributed occurrences of floods. A POT sample is created using all values exceeding a predefined threshold. The main advantage of the POT approach is therefore the consideration of all severe floods within a flood intensive year, while years with no extreme events are neglected. Thus, a POT time series captures more information concerning the entire flood characteristics of a river than using AMF. The key challenge of the POT approach, however, is the threshold selection, since statistical methods

(e.g. extreme value distribution) may react very sensitive to different thresholds. Selecting suitable thresholds is therefore a complex task representing the main difficulty associated with the POT approach (Lang *et al.*, 1999). Additionally, the independence of the individual events has to be assured as well.

Lang *et al.* (1999) reviewed some threshold selection techniques. An important factor in the threshold selection is the mean number N of events per year. They recommend that there should be at least a mean number of floods of N = 2 or 3 per year. A common threshold selection criteria is to use a standard frequency factor f, so that the threshold can be estimated from the daily mean discharge series Q by:

$$u=\mu_Q+f\cdot\sigma_Q \quad (1)$$

where μ_Q and σ_Q are the mean and standard deviation of the daily mean discharge series Q, respectively. Thus we use frequency factors, so that we consider 2, 3, and 4 events per year.

To analyse whether the 100-years flood event is affected by the latest extreme events, we use a quasi-nonstationary extreme value approach based on both the Generalized Extreme Value distribution (GEV) and the Generalized Pareto Distribution (GPD). We start with analyzing a time window from the beginning of the time series to 1950. The time series is then constantly extended by one year. For each time series the 100-year return level is calculated.

Besides analysing the 100-year level the 2013 event at gauge Neu Darchau was assessed by using record theory, since this event is a new record (i.e. a value higher -or lower- than any previous value in the data). Following Rahmstorf & Coumo (2011) we like to investigate, whether this record could be assessed as a result of natural variability or rather of climate change.

4 RESULTS

4.1 Changes in the 100-year flood event

The time dependent behavior of the 100-year event at gauge Neu Darchau is shown in Figure 2. The black thick lines represent the 100-year event, whereas the red lines show the 95 % confidence intervals. It can be stated that the results from the r-largest time series (AMF, AMFr2, AMRr3) have a larger variability than the POT based results. The results indicate that the 100-year level decreases from 1950 to about 1980 and slightly increases until 2013 or remains constantly. In summary, it can be stated, using the here applied quasi non-stationary extreme value approach, that the time dependent behaviour of return discharges of the main flood indicators is not significantly affected by the extreme events in 2002, 2006, and 2011. However, it can also be observed, that the record value in 2013 has a small influence on the 100-year event.

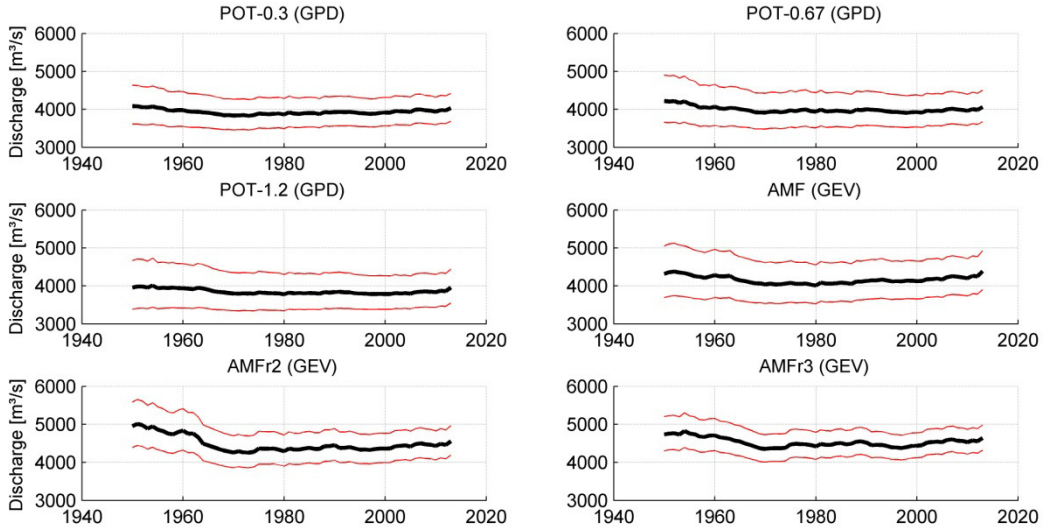


Figure 2: Time-dependent 100-yr return levels of discharge (thick black line) at gauge Neu Darchau from 1950 to 2013 using different flood indicators derived by extreme value statistics. The red lines refer to the 95 % confidence levels

The same results from the gauge Dresden are shown in Figure 3. Also the 100-year event is steadily decreasing from 1950 until 2001. In 2002 a positive shift is identifiable and the same is true for the year 2013. Thus, in summary it can be stated that the 100-year flood event at gauge Dresden is much more affected by the extreme events in 2002 and 2013 than at gauge Neu Darchau.

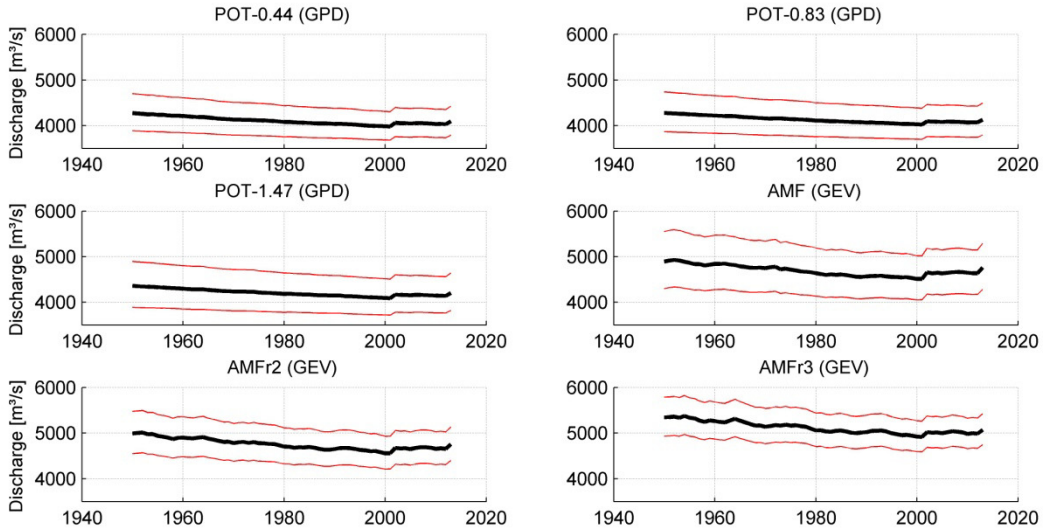


Figure 3: Time-dependent 100-yr return levels of discharge (thick black line) at gauge Dresden from 1950 to 2013 using different flood indicators derived by extreme value statistics. The red lines refer to the 95% confidence levels

Figures 4 and 5 contain the daily mean discharge time series of the gauges Neu Darchau and Dresden with different return levels derived from the GPD analyses with 2 events per year in mean, i.e. POT-0.67 for Neu Darchau and POT-0.83 for Dresden. It is worthwhile to mention, that the return levels at both gauges are in the same magnitude (e.g. $HQ_{100,NeuDarchau}$ = 4060 m³/s and $HQ_{100,Dresden}$ = 4127 m³/s), although the gauge Neu Darchau is approximately 500 km downstream of the gauge Dresden. This is a result of the wave transformation processes along the river. From Figure 4 it is visible, that until 2012 the highest flood event was in the range of a 50-year return level and the record value in 2013 can be assessed as a 100-year flood event. The gauge Dresden exhibits a different system behavior. The 100-year return level was exceeded 5 times within the observed time series. The flood event from June 2013 can be ranked as a 100- to 200-year flood event and is the fifth largest flood event.

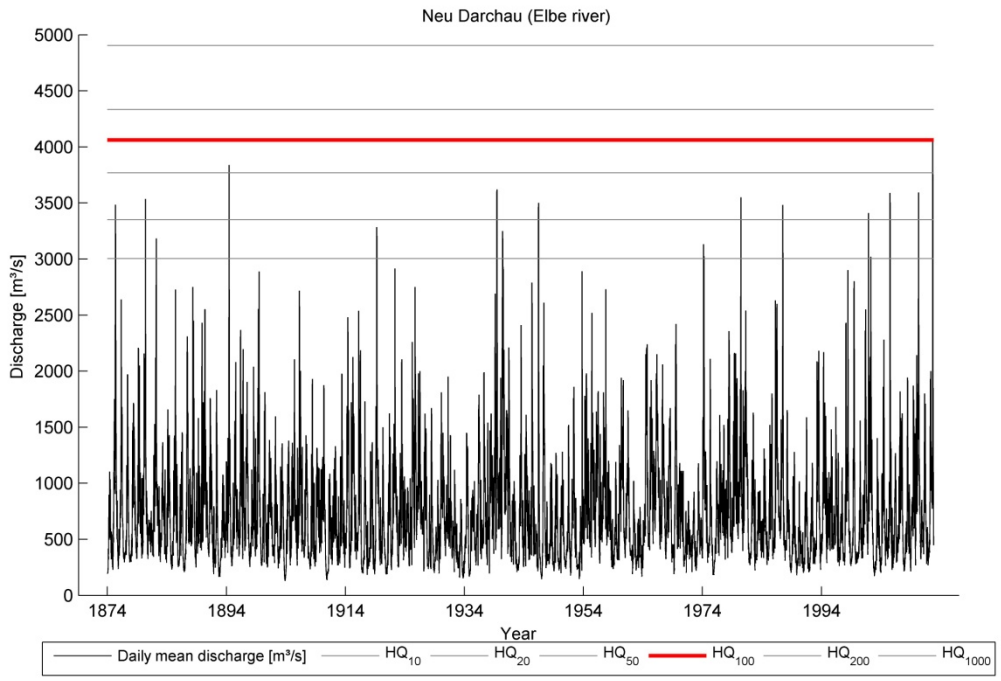


Figure 4: Time series of daily mean discharge of gauge Neu Darchau with different return levels derived from GPD analyses with 2 events per year in mean

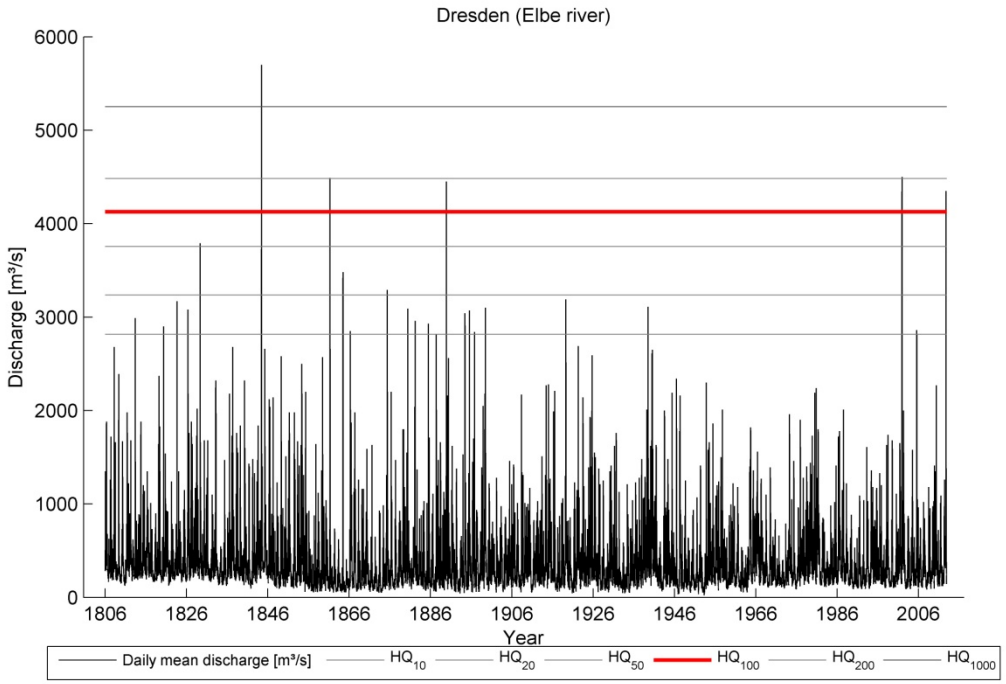


Figure 5: Time series of daily mean discharge of gauge Dresden with different return levels derived from GPD analyses with 2 events per year in mean

4.2 Record theory

Since the flood event in 2013 at the gauge Neu Darchau is a new record, it may be interesting to investigate, whether this record is linked to climate change or belongs to the natural variability of the river system. This can be achieved using the record theory. Records are values which are higher than any previous values in the dataset. Since the first value of a dataset is per definition a record, the times series of the gauge Neu Darchau contains 6 records, based on the annual maxima floods from 1884 to 2013 (Figure 6).

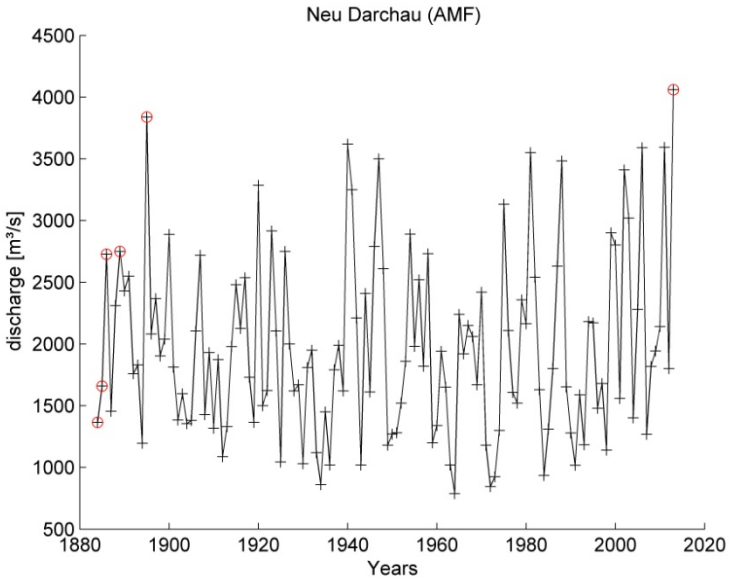


Figure 6: Time series of annual maxima of the gauge Neu Darchau from 1884 to 2013 and identified records (circles)

In a stationary climate the number of records declines as 1/n, so the expected number in a time series of 100 years length is 5.2 (Rahmstorf & Comou, 2011). Giving an example: In a time series of 100 years length, the first value is a record with the chance 1/1, the second with the chance 1/2 and so on. The 100th value is a record with the chance 1/100. Summing up all the probabilities result in an expectation value of 5.2 records for a 100-year time series (Figure 7). In a time series with a length of 130 years (as investigated here) this number increases to 5.4. From this analysis the observed number of records (6) seems to be proper.

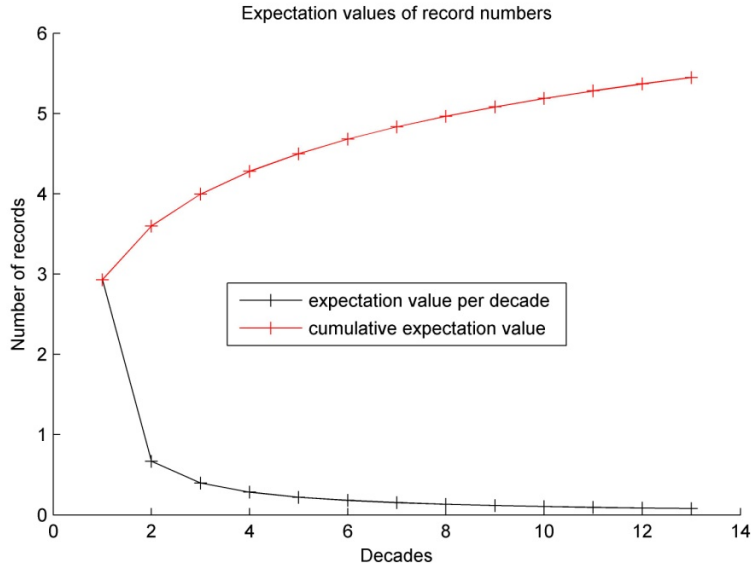


Figure 7: Expected number of records per decade of a stationary time series. The number of records per year declines with 1/n

In reality the time series of the gauge Neu Darchau can not be assessed as a full stationary system. With any kind of non-stationarities the expected number of records within a year or decade will change. To analyse this more in detail we fitted a non-linear trend model (LOWESS filter) to the time series of the annual maxima of the gauge Neu Darchau. By using Monte-Carlo-simulations the expected number of records per decade can be calculated in a non-stationary system (Figure 8). From this we derive, that there will be an expected number of records for the last decade for the gauge Neu Darchau of 0.46. The respective number for the stationary climate is 0.08, determined by the 1/n law. Since in the last decade a record was observed (2013 flood event) there is a 83 % [(0.46 – 0.08)/0.46] chance, that this record is due to a non-stationary system behavior (e.g. climate change).

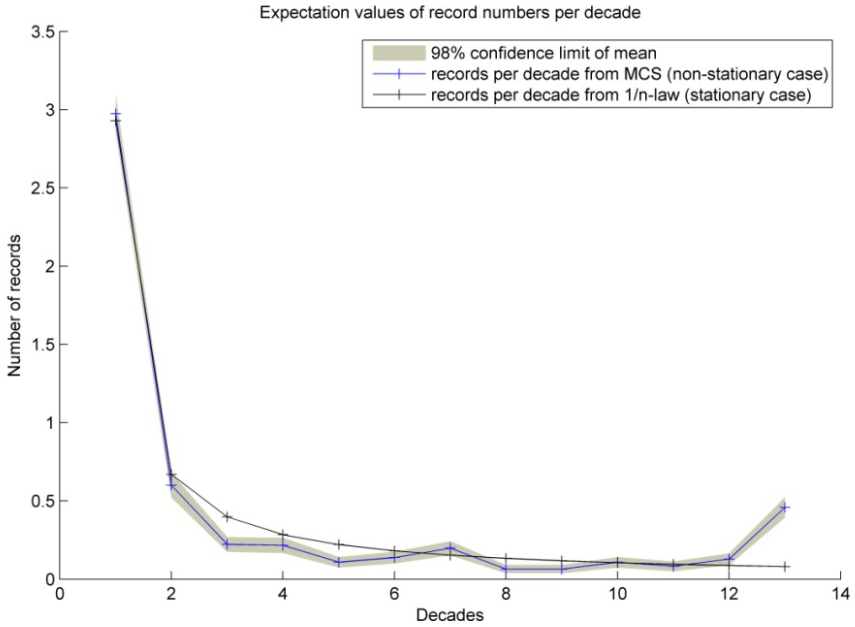


Figure 8: Comparison of the expected number of records per decade in a stationary climate and a non-stationary climate for the gauge Neu Darchau

5 DISCUSSION AND CONCLUSIONS

In this study we investigated, how the 100-year level at the gauges Neu Darchau and Dresden is affected by the latest extreme events in the years 2002, 2006, 2011, and 2013. It was shown, that the 100-year flood event at the gauge Neu Darchau is less affected by the extreme events than at the gauge Dresden. Additionally, from record theory the result was obtained, that there is a 83 % chance, that the 2013 record is rather due to non-stationary changes in the system than due to the natural variability.

To answer the question, whether the 100-year flood event serves as a reliable basis for future flood protection design purposes, a detailed cost-benefit-analysis should be conducted. Since in this study we only focused on time-dependent changes in the 100-year return level, a final conclusion of the further applicability of the 100-year flood event can not be made. But the results indicate, that the 100-year return level was exceeded several time within the observation period (only true for Dresden) and that the extreme events in the last decades or at least the 2013 event will lead to higher 100-year levels at both gauges. However, in the context of the European legislation (flood directive), the public perception of the latest extreme events and the damages caused by the latest extreme events it seems to be worthwhile to think about higher flood protection levels, e.g. towards a 200-year or 500-year event.

6 REFERENCES

Acero, F.J., Garcia, J.A. and Cruz Gallego, M. (2011): *Peaks-over threshold study of trends in extreme rainfall over the Iberian*, Peninsula, J. Climate 24, pp.1089–1105.

Bayliss, A.C. and Jones, R.C. (1993): *Peaks-over-threshold flood database: Summary statistics and seasonality*. Report No. 121. Institute of Hydrology: Wallingford.

Coles, S. (2001): *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag: New-York.

FEMA (2013): http://www.fema.gov/media-library-data/20130726-1545-20490-9247/frm_acts.pdf

Lang, M. Ouarda, T .B.M.J. and Bobée, B. (1999): *Towards operational guidelines for over-threshold modeling.* Journal of Hydrology 225, pp.103–117. doi:10.1016/S0022-1694(99)00167-5.

Leadbetter, M.R. (1991): *On a basis for 'Peaks over Threshold modeling*. Statistics & Probability Letters 12, pp.357–362.

Rahmstorf, S. and Coumou, D. (2011): *Increase of extreme events in a warming world*, PNAS, 44,108.

Rölver, N. (2012): *Discharge on 24/25 March 1888 at gauge Neu Darchau*. WSA Lauenburg. pers. Comm.

Svensson, C., W., Kundzewicz, Z. and Maurer, T. (2005): *Trend detection in river flow series: 2. Flood and low-flow index series*. Hydrological Sciences Journal, 50 (5), 811-824.

WSA Lauenburg (2012): http://www.wsa-lauenburg.wsv.de, official website of the gauge operator.

Effect of ENSO-based climate variability in the estimation of flood events in Argentina

Ana Claudia Callau Poduje[1] and Rafael Seoane[2]

[1]Institute of Water Resources Management, Hydrology and Agricultural Hydraulic Engineering, Leibniz University Hannover, Hannover, Germany, Email: callau@iww.uni-hannover.de

[2]University of Buenos Aires (UBA), National Water Institute (INA), National Research Council (CONICET), Buenos Aires, Argentina

Abstract

The influence of the El Niño-Southern Oscillation (ENSO) phenomenon depends on the region, and may be significant in some areas. ENSO-based variations may result in different hydrological responses according to the region. The aim of this work is to study the effect of the ENSO in the flood characteristics for different rivers in Argentina. Flood events are characterized by peak flow and volume and the dependence structure between these two variables is modeled with copulas. A set of large catchments belonging to different climatic regions in Argentina is included in the analysis. Flood events are separated according to the ENSO index in the three climatic states: cold, warm and neutral.
A copula model is defined for each river without accounting the climate state information, i.e. including the total set of floods. Then additional copulas are used to model the dependence structures between the variables belonging to the different climatic states.
Considering the ENSO effect results in models to represent the floods that differ from the ones obtained including all the events. The difference between the models with and without the account of the climate state varies according to the climatic region. These results suggest that the relationship between the hydrological response of large catchments in Argentina and the ENSO-based climate variability varies along the country and depends on the climatic region.

1 INTRODUCTION

Several existing studies indicate that the ENSO is a significant factor in determining inter-annual rainfall variability in South America (Prieto, 2007). In Argentina the impact of ENSO phenomenon varies according to the region and to the phase, i.e. warm (El Niño) or cold (La Niña). During El Niño events the Andean mountain chain (western limit of Argentina) registers many intense snow storms and higher precipitation than the normal values in winter, whereas strong and intense storms are observed in the north-east of the country during the summer and autumn. The Northeastern region of Argentina is one of the most exposed areas in South America to flood-related risk disasters (Prieto, 2007). Therefore this region has been studied by several authors in order to define the relationship between the phases of ENSO and the major discharge anomalies. Camilloni and Barros (2003) found a strong relationship between the ENSO and the high flows registered in the Upper Paraná river, whereas the relationship associated to monthly or seasonal discharges is weak. Their results show that about two thirds of the major discharge anomalies occurred during El Niño events and none of these major anomalies occurred during La Niña phase. Most of the existing studies focus on the analysis of the relationship between flows and ENSO events. Chiew & McMahon (2002) studied the global ENSO-streamflow teleconnection for different catchments around the word, and concluded that the correlation between ENSO and streamflow can be used for forecasting flows. Kiem *et al.* (2003) studied the flood risk across New South Wales in Australia, and state that the La Niña events are dominant drivers of elevated flood risk.

The aim of this work is to study the effect of the ENSO in the flood characteristics for different rivers in Argentina. Flood events are characterized by peak flow and volume along with their dependence structure. This work is meant to be a first approach to understand whether it is important to consider the ENSO phenomenon in the design of hydraulic structures in different areas of the country.

2 METHODOLOGY

2.1 Definition of flood characteristics and ENSO events

The definition of the variables characterizing the flood consists of the following steps of calculation:

- Separation of base flow and direct flow,
- Definition of the flood period, in terms of maximum annual/seasonal flow,
- Identification of start and end dates of the flood,
- Estimation of the variables describing the flood event: volumes (total, direct and base), peak flows (total, direct, base), duration and time to peak.

The flood characteristics used in this work are the total peak flow and total volume. The dependence structure describing the joint behavior of these two variables is included in the study. The rest of the variables are not analyzed here, but are going to be studied in future works.

The ENSO phenomenon is characterized by episodic warming or cooling of the ocean temperatures over and above the seasonal cycle in the Equatorial Pacific. This oscillation of the ocean-atmosphere system has important consequences in the weather not only in the area of its occurrence, but even at far away locations, being the most notable and pronounced example of interannual global climate variability. In case of warming of the sea surface temperatures (El Niño phase) the consequences are heavy rainfall over the coastal regions of northern Perú and Ecuador, but severe droughts in eastern Australia. On the other hand, the cooling of the sea surface (La Niña) results in droughts in South America and heavy rainfall and even floods in eastern Australia (Rakhech & Singh, 2009).

To study the relationship between the extreme ENSO phases and the characteristics of the flood events, the months in which the flooding occur are identified as warm, neutral or cold ENSO phases. This procedure allows grouping the flood events into warm and cold phases and comparison of both groups with the complete series. The seasons and years corresponding to each ENSO phase are defined according to the data provided by the National Oceanic and Atmospheric Administration (NOAA) which can be accessed in their website http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml (accessed August, 2013). This data base includes a list indicating warm and cold events on a three month time basis since 1950 to the present. In order to confirm and extend the data base to periods previous the year 1950, some publications were used, which include longer time series of ENSO phases (see Kiladis & Diaz, 1989; Trenberth, 1997). The result is a classification of the years as warm, neutral or cold ENSO phase starting in 1900.

2.2 Copula theory and joint return period

In this work a bivariate statistical frequency analysis of the variables describing the floods is carried out using copulas. A copula, denoted as $C(u_1,u_2,\ldots,u_n)$, is a function that enables modeling the dependency structure of random variables, independently of their marginal distributions. The link between the copula and the multivariate distribution is provided by the theorem of Sklar (see Nelsen, 2006) with the following equation:

$$F_{X1,X2,\ldots,Xn}(x_1,x_2,\ldots,x_n) = C[F_{X1}(x_1), F_{X2}(x_2),\ldots,F_{Xn}(x_{n1})] \quad (1)$$

where $F_{X1,X2,\ldots,Xn}(x_1,x_2,\ldots, x_n)$ is the joint cumulative distribution function with the continuous marginal distribution functions of the random variables: $F_{X1}(x_1)$, $F_{X2}(x_2),\ldots,F_{Xn}(x_n)$. When the models representing each marginal distribution, i.e. function and parameter(s), and the copula function along with it(s) parameter(s) are chosen the model representing the joint behavior of the random variables is defined. The univariate models considered for fitting the series of peak flow and volume and defining the marginal distributions are: Gamma, Generalized Extreme Value (GEV), Gumbel, Log-Normal (2 and 3 parameters), Normal (2 and 3 parameters), Pearson (3 parameters and Weibull (2 and 3 parameters). The methods of estimating the parameters include method of moments, L-moments and the maximum likelihood method. The goodness of fit of the different models is assessed using the Kolmogorov-Smirnov (KS) and Cramer-Von-Misses (CvM) criteria, and by a visual inspection of the quantile plots.

The relationship between flood peak and volume is studied considering the following Copula

functions Clayton, Frank, Galambos and Gumbel, which are described by one parameter. The methods used for estimating the parameters are the moment-like and pseudo-likelihood methods, which are rank-based and rely completely on the relative ranks of joint variates (for further details refer to Genest & Favre, 2007). The copulas considered in this work are able to model the whole range of positive dependence between variables. In order to compare the empirical and parametric copula probabilities, a CvM type of test is applied (Genest & Favre, 2007; Chowdhary *et al.*, 2011). Large values of this statistic can lead to the rejection of the model under consideration. The analysis and synthesis of marginal and joint distributions presented in this work have been performed using R (R Development Core Team, 2012) a free software environment for statistical computing. The packages used are the Fitdistrplus package (see Delignette-Muller *et al.*, 2010), the Lmomco package (see Asquith, 2010) and the Copula package (see Yan, 2007; Kojadinovic & Yan, 2010; Hofert & Maechler, 2011).

A common criterion employed to design hydraulic structures is the return period (*T*) of the extreme hydrological events. The univariate case considers only one variable as critical according to the equation 2, whereas the bivariate joint return period includes two variables. In the bivariate case, an event can be defined as critical if either one of the two random variables exceeds a given threshold ("OR" case) or if both variables are larger than some prescribed values ("AND" case). These two types of joint return periods depend on the joint cumulative distribution function, or Copula, according to the following equations:

$$E\left(T_{\mathrm{X}}\right)=\frac{E\left(L\right)}{P\left(X\geq X_{\mathrm{T}}\right)}=\frac{E\left(L\right)}{1-P(X<X_{\mathrm{T}})} \tag{2}$$

$$T_{\text{"OR"}}=E\left(T_{\mathrm{XY_{or}}}\right)=\frac{E\left(L\right)}{P\left(X\geq X_{\mathrm{T}}\text{ or }Y\geq Y_{\mathrm{T}}\right)}=\frac{E\left(L\right)}{1-F(X_{\mathrm{T}},Y_{\mathrm{T}})} \tag{3}$$

$$T_{\text{"AND"}}=E\left(T_{\mathrm{XY_{and}}}\right)=\frac{E\left(L\right)}{P\left(X\geq X_{\mathrm{T}}\text{ and }Y\geq Y_{\mathrm{T}}\right)}=\frac{E\left(L\right)}{1-F\left(X_{\mathrm{T}}\right)-F\left(Y_{\mathrm{T}}\right)+F(X_{\mathrm{T}},Y_{\mathrm{T}})} \tag{4}$$

where *E(L)* is the expected interarrival time, under the assumption that this variable is described by an identical and independent distribution (see Shiau, 2003). When the equations are applied to annual maximum series, *E(L)* equals to1 year due to the fact that one event is selected for each year. Figure 1 illustrates the different concepts of return period in terms of probabilities and values of the variables. As an example the maximum annual values associated to a return period of 100 years, i.e. a probability of 0.99, are presented. The joint return periods can be presented by contours of equal probability whereas for the univariate case only one pair of values is identified (see upper right image in Fig. 1). These probabilities are transformed into values of peak flow and volume through the corresponding marginal distributions (see upper left and lower right images in Fig. 1). The contours corresponding to a return period of 100 years, both for the "OR" and "AND" cases along with the value corresponding to the univariate case are shown in the lower left image (see Fig. 1). The plotted values, which are only illustrative, correspond to a Clayton copula and the marginal distributions are General Extreme Value for the peak flow series and Weibull with 3 parameters for the volume series.

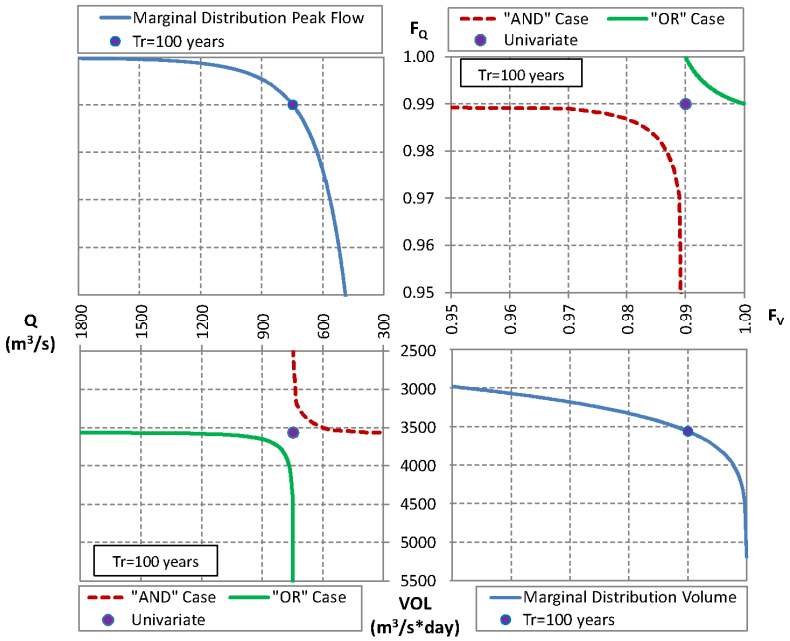


Figure 1: Comparison between classical and joint return periods, in terms of probabilities and their corresponding values of the variables

3 DATA

The proposed procedure was applied to daily streamflow data registered in four different rivers in Argentina. The selected stations and rivers include more than 95 complete years to study the flood events. The chosen rivers belong to different climatic regions in the country, and thus may respond to the ENSO phenomenon in different ways. These rivers are: Paraguay, Paraná, Neuquén and Limay, and some basic information about the stations can be found in the following table.

Table 1: Description of the gauging stations and series used for this study

River	***Station***	***Catchment Area* (km²)**	***Mean Flow* (m³/s)**	***Latitud* (ºS)**	***Longitud* (ºS)**	***Altitud* (m a.m.s.l.)**
Paraná	Corrientes	1950000	17102	27.48	58.83	60
Paraguay	Puerto Bermejo	1100000	3770	26.93	58.51	50
Neuquén	Paso de los Indios	30843	312	38.53	69.41	498
Limay	Paso Limay	26400	736	40.53	70.43	538

The location of each gauge station within the country as well as the zonification of the water regime is shown in the following figure (see Fig. 2). This zonification map was developed by Moyano & Diaz (2006) based on the climatic characteristics and indicates the water regime as a function of the aridity index, which is a value that relates the mean annual precipitation over the evaporation (see UNEP, 1997). The selected rives belong to very different climatic regions, the Paraná and Paraguay rivers belong to climates which are “hidric” to “hiper hidric”, indicating enough rain all over the year compared to the ecaporation, whereas the catchments belonging to the Limay and Neuquén rivers are located in zones classified in most of the surface as “arid” and “semi-arid”.

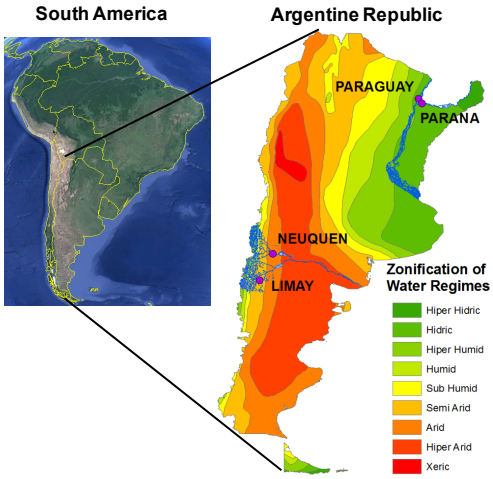


Figure 2: General location map and Argentine Republic map with zonification of water regimes and gauge stations used in this work (marked with purple circles)

The Paraná and Paraguay rivers belong to the La Plata system, which registers more than 85% of the total mean flow in the whole country. Therefore most of the population, urban development and economic activities are concentrated in this area. The Paraná river has an important hydroelectric project which generates 15 % of the energy consumed in the country, whereas the Paraguay river is an important navigation corridor that connects Bolivia and Paraguay and the northern part of Argentina to the Atlantic ocean. These rivers have one flooding period due to the rainy season which is in the summer (December to March). The Paraná river presents the flooding period by the end of the summer, whereas the flooding in the Paraguay river is delayed to the autumn months due to the regulation effect of the Pantanal wetlands which feed the Paraguay river. Camilloni & Barros (2003) state that the maximum discharges of the Paraná River basin are due mainly to the precipitation in the Upper Paraná, which is upstreams from the confluence with the Paraguay river, while the contribution from the Middle and Lower Paraná basins to extreme discharges is relatively small.

Besides de La Plata system, the most important rivers in Argentina are the ones which have their springs in Cordillera de Los Andes and flow westwards to the Atlantic Ocean, through very arid and semi-arid areas without any important tributaries on the way. The rivers Neuquén and Limay belong to the Río Negro system, which is an important system due to the fact that is presents most of the flood control projects as well as hydropower plants, representing 43 % of the total hydropower generation of the country (Secretaría de Energía, 2004). These rivers have two types of flooding seasons: one in winter due to the rain and a very different one in summer due to the melt of the snow which might be combined with some rain events.

4 RESULTS AND DISCUSSION

The floods corresponding to each season/year were identified for each of the four rivers. After distinguishing the floods and the variables characterizing them, the different events were grouped according to the ENSO-phase they belong. The variables characterizing the floods are studied for the whole record period and for the groups corresponding to the warm and cold phases. The subsequent paragraphs and figures are presented in order to understand the joint behavior of the variables and to describe the magnitude of the series associated to the different ENSO phases. The following figures show the flood characteristics observed in the Paraná (Fig. 3) and Paraguay (Fig. 4) rivers.

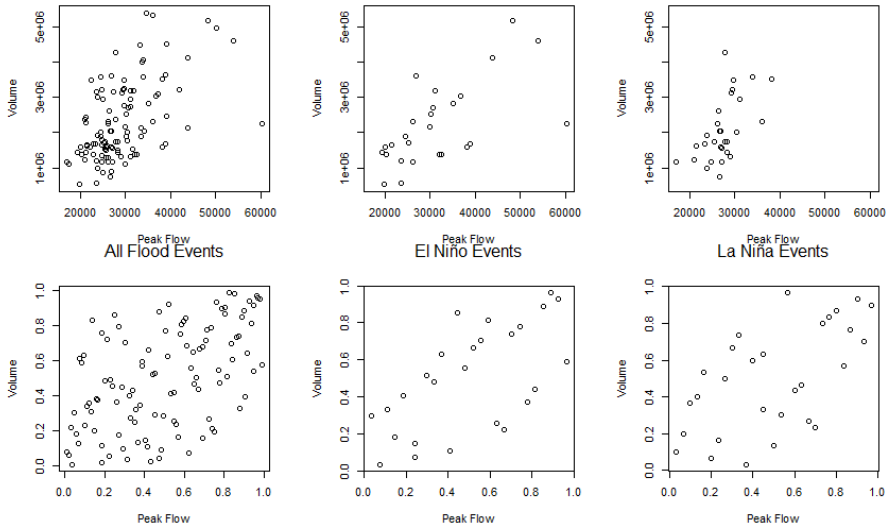


Figure 3: Pairs of variables (first row) and pseudo-variables (second row) describing the floods registered in the Paraná river

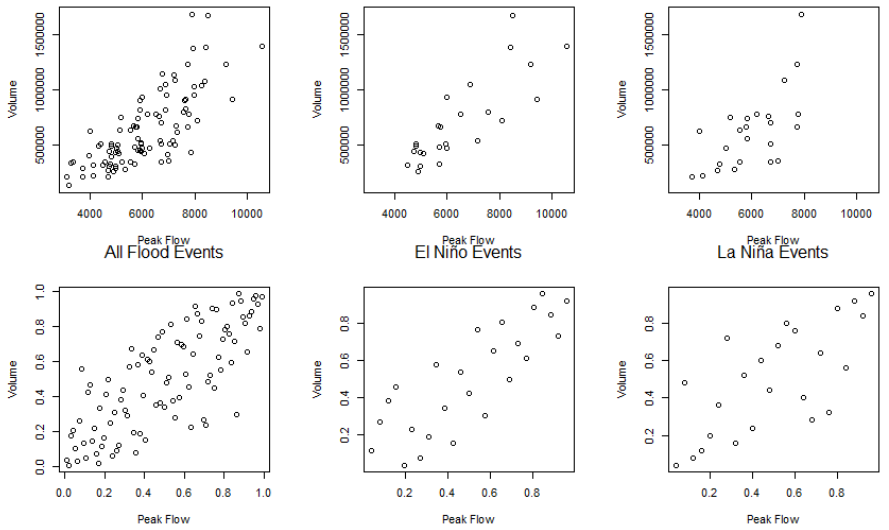


Figure 4: Pairs of variables (first row) and pseudo-variables (second row) describing the floods registered in the Paraguay river

Figures 3 and 4 indicate that the peak flow values observed in the Paraná and Paraguay rivers reach lower magnitudes and are less variable during La Niña phase. For the volume series the Paraná river shows a similar behavior, whereas in the Paraguay river the median of the floods observed during La Niña phase are higher. Some extreme events are observed in the plots of All Flood Events, which indicate floods with high values of peak flow and volume that were not registered during any ENSO phase, they are more frequent in the Paraná river. The correlation between the variables describing the flood events are studied for the cold and warm phases of the ENSO in order to decide if any of these phases has some effect on this joint behavior. The results are shown in Figure 5 (left graph).

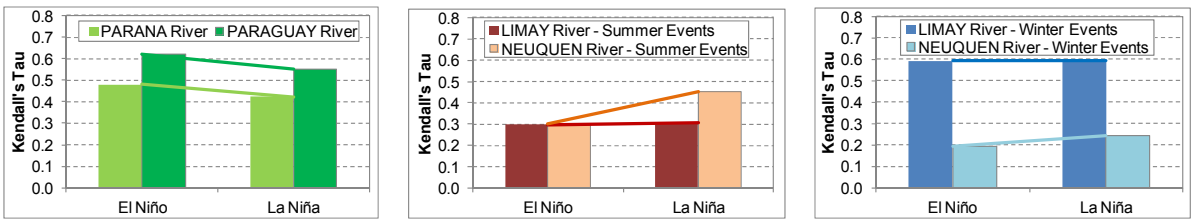


Figure 5: Correlation between the variables describing the floods observed in the Paraná and Paraguay (left), summer (middle) and winter (right) in the Limay and Neuquén rivers

Both rivers, Paraná and Paraguay, show that the correlation between the variables peak flow and volume are higher for the events associated to El Niño phenomena. These results can be observed in the figures showing the pseudo-variables (see second rows of Fig. 3 and 4), in which the pairs of values corresponding to El Niño events are more concentrated in the main diagonal, compared to the La Niña and All Events pairs of values. These behavior indicates that in case of El Niño it is more likely to have events in which for a high peak flow a high value of volume would be observed, whereas for La Niña phases the variables characterizing the floods are not so strongly correlated so a high peak flow can be associated to a low volume, or vice versa. The correlation between the variables describing the winter and summer flood events observed in the Limay and Neuquén rivers are shown in the following graphs (see Fig. 5 middle graph). The winter floods registered in the Limay and Neuquén rivers are a consequence of the rainfall in the region during the winter period. The correlation parameters indicate that for the Limay river, the floods corresponding to El Niño and La Niña have a similar association so there is no clear difference between the flood belonging to both phases. In the Neuquén case the floods observed during La Niña phase show a higher correlation compared to the ones observed during El Niño phase, though the difference is not so high (see Fig. 5 right graph). The next figures show the winter flood characteristics observed in the Limay (Fig. 6) and Neuquén (Fig. 7) rivers.

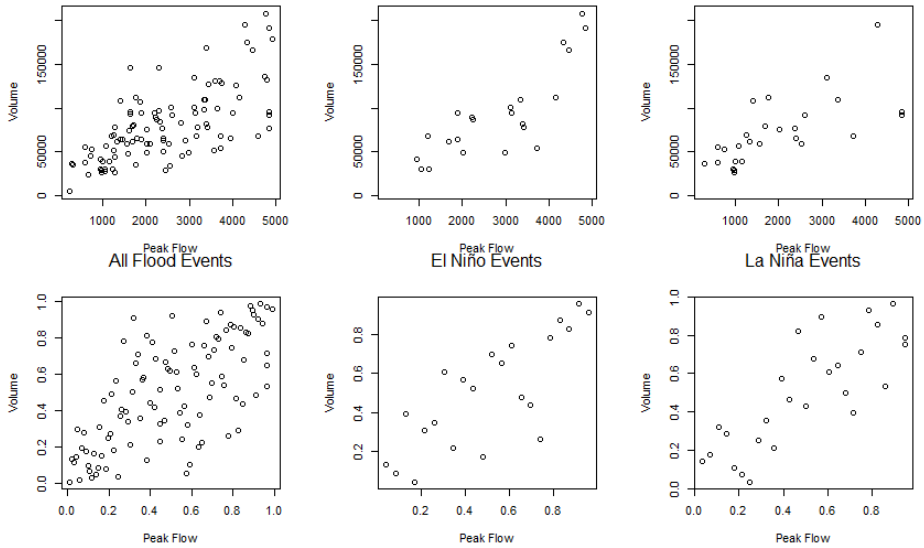


Figure 6: Pairs of variables (first row) and pseudo-variables (second row) describing the winter floods registered in the Limay river

Figure 6 shows that the Limay's winter floods observed during El Niño present peak flows and volumes that are more concentrated towards high values, but if all the events are considered some additional extreme cases are included in the plot and some of these events were not observed in El Niño years. Figure 7 indicates that both peak flow and volume series observed in the Neuquén river reach lower levels and are less variable during La Niña phase. The total serie of flood events shows a higher concentration of points in the lower left corner, which indicates events with low values of peak flow and volume, which might not be critical for the hydraulic structures.

The summer floods registered in the Limay and Neuquén rivers are consequences of the snow melt due to the high temperatures, which can be combined with some rain events. The results for this case are similar than for the winter floods, i.e. the Limay river does not show a difference when the foods belonging to the two phases are compared, whereas the Neuquén river shows a higher correlation between the variables characterizing the La Niña phase (see Fig. 5 right graph). The following figures show the summer flood characteristics observed in the Limay (Fig. 8) and Neuquén (Fig. 9).

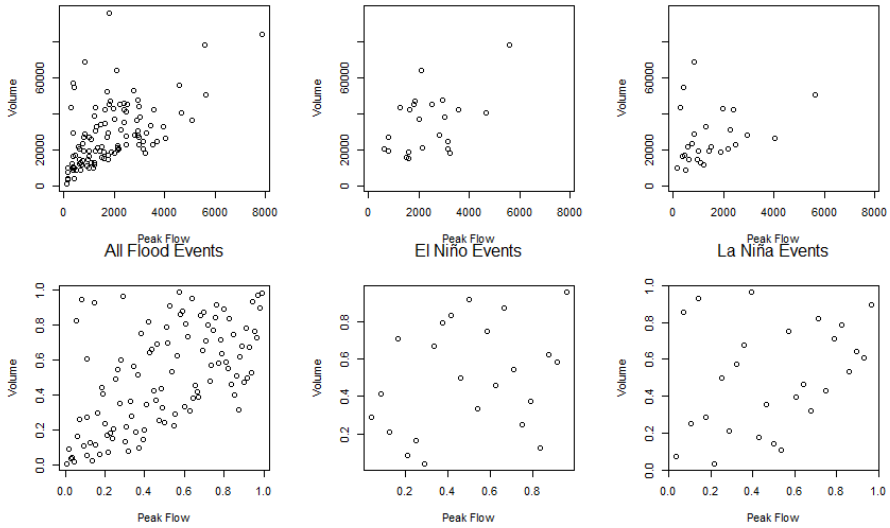


Figure 7: Pairs of variables (first row) and pseudo-variables (second row) describing the winter floods registered in the Neuquén river

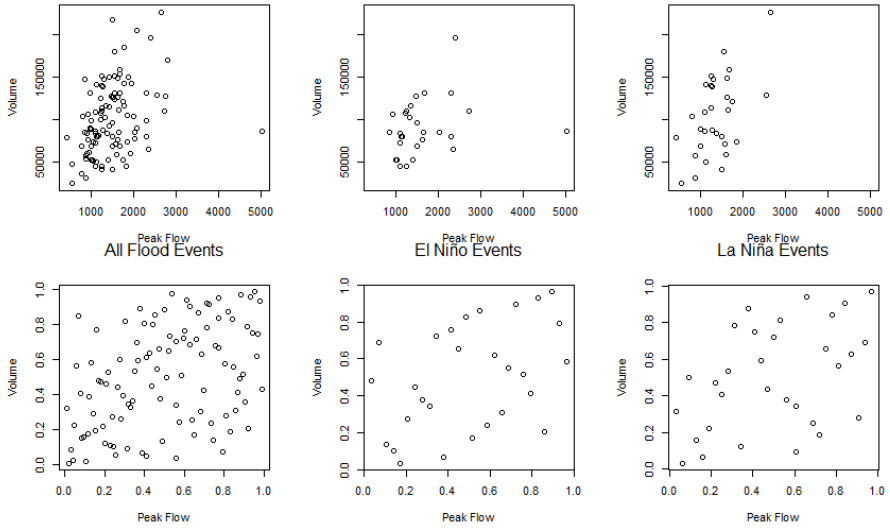


Figure 8: Pairs of variables (first row) and pseudo-variables (second row) describing the summer floods registered in the Limay river

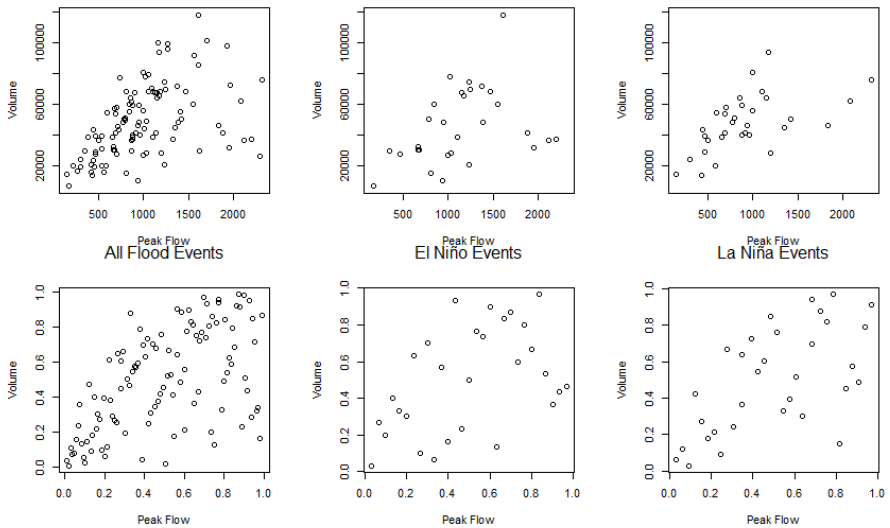


Figure 9: Pairs of variables (first row) and pseudo-variables (second row) describing the summer floods registered in the Neuquén river

Figure 9 shows that the floods observed during La Niña present lower and less variable peak values but higher and more variable volume values. If all the events are considered, the plot shows a high concentration of events in the region of low peak flows, although for the volume values the distribution is more variable. Figure 10 indicates that the peak flows and volumes are less variable for events associated to La Niña. In case of considering all flood events, more extreme cases are included in the analysis, that show high values of peak flow associated to high volumes.

In the previous paragraphs two concepts were discussed regarding the flood events associated to the ENSO phases: the joint behavior of the variables describing the events and the magnitude of the observed variables. It is the aim now to analyze what is the effect of combining these two concepts in terms of a design criterion. For this purpose the different groups of flood events are studied and the combination of variables associated to a joint return period of 100 years (OR case) are estimated. Some of the results are presented in the subsequent paragraphs and figures. Each of the rivers was studied separately, and the aim was to find the most suitable copula model to represent the joint behavior of the variables describing the floods grouped in the different cases (El Niño, La Niña and all events together). Different copulas were considered and the one that was not rejected with a 5% of confidence level for the three cases was selected. The marginal functions were also chosen so that one type of function appropriately represents the three sets of variables. This procedure was followed in order to avoid the effect of the copula model and marginal functions in the results and to focus on the use of the different series in the analysis. The copula used to model the joint behavior of the variables describing the floods observed in the Paraná river is the Galambos, and the parameter is estimated using the method of maximum pseudo-likelihood. The probability distribution describing the marginal behavior of the peak flow is the Log-Normal (3 parameters) whereas for the volume the best representation is performed by the Pearson (3 parameters). The following images (see Fig. 10) show the possible combinations of pseudo-variables and variables associated to a 100 years return period.

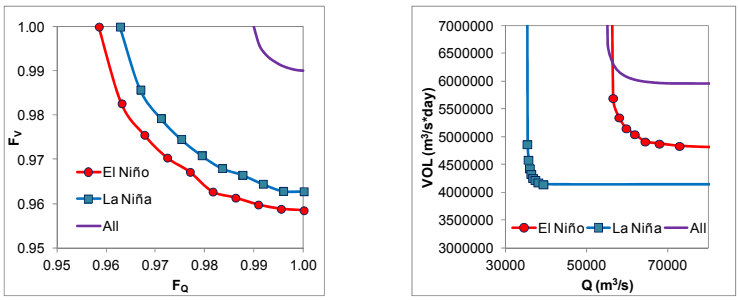


Figure 10: Pairs of pseudo-variables (left) and variables (right) associated to a joint return period of 100 years corresponding to the Paraná river

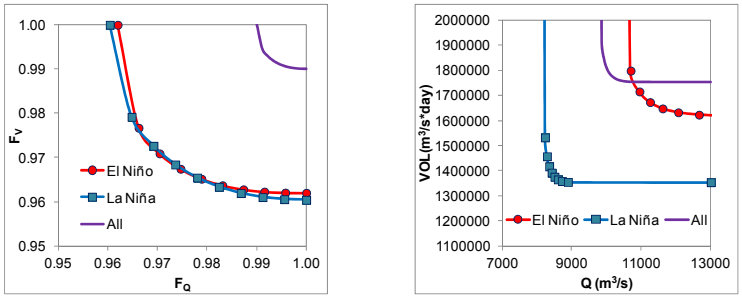


Figure 11: Pairs of pseudo-variables (left) and variables (right) associated to a joint return period of 100 years corresponding to the Paraguay river

The curves representing the pseudo-observations vary due to the fact that the phenomena El Niño and La Niña are not observed every year but every approximately 4 years, therefore the expected inter-arrival times (see Eq. 2 to 4) are higher and lower than 4, for El Niño and La Niña sub-series respectively. The floods registered in the Paraguay river are described by a Gumbel copula, and the parameters are fitted with the method of moments. The peak flow series are represented by the Generalized Extreme Value distribution, whereas the volume by the Pearson (3 parameters). Figures 10 and 11 show that some of the pairs of values associated to a 100 years return period are more critical for the El Niño phase. If the peak flow is an important design criteria for the structure

then, if the design focuses on the El Niño events, the design parameters will be higher. This conclusion is especially important for the Paraguay river in which the difference between considering the El Niño events and considering all events for the design results in values of 9800 versus 10700, both associated to high volumes (see Fig. 11 upper right corner of the second graph). On the other hand if only the volume is important for the structure under design, then considering all the events will result in more critical values. The same analysis was performed for the summer and winter flood events observed in the Limay and Neuquén rivers. In case of the winter events, the floods observed in the Limay were fitted by the method of maximum pseudo-likelihood to a Gumbel copula, with marginal distributions Weibull (3 parameters) for the peak flows and Generalized Extreme Value for the volume. The floods observed in the Neuquén are represented by a Clayton copula (method of maximum pseudo-likelihood) with a Pearson (3 parameters) distribution for the peak flow and volume series. The resulting curves representing the Limay winter floods indicate that considering all events results in more critical design parameters. The Neuquén winter floods corresponding to El Niño show that the events with a 100 years of return period are more critical in terms of volumes, which for high values of peak flow are 85000 m^3/s*day versus 78800 m^3/s*day if all the events are considered (see Fig. 12 left image).

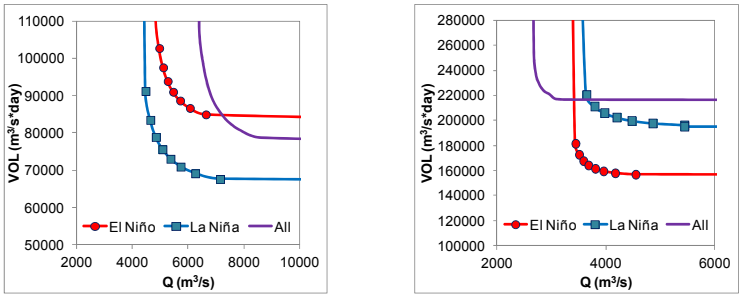


Figure 12: Pairs of variables corresponding to the winter floods of the Neuquén river (left) and to the summer floods of the Limay river (right) associated to a joint return period of 100 years

In case of the summer flood events, the Limay events were fitted by the method of maximum pseudo-likelihood to a Galambos copula, with marginal distributions Log-Normal (3 parameters) for the peak flows and Weibull (3 parameters) for the volume. The floods observed in the Neuquén are represented by a Clayton copula (method of maximum pseudo-likelihood) with a Generalized Extreme Value distribution for the peak flow series and a Pearson (3 parameters) for the volume series. The Limay summer floods corresponding to both La Niña and El Niño show that the events associated to 100 years of return period are more critical in terms of peak flows (see Fig. 12 right image). For the Neuquén summer floods considering all events results in more critical design parameters.

5 CONCLUSIONS

The effect of the ENSO phenomenon in the estimation of flood events in Argentina was performed focusing on two particular issues: first the description of the flood events in terms of variables characterizing them, and in a second step a joint return period analysis was performed in order to decide whether it is meaningful to focus on one phase of the ENSO phenomenon that could be more critical in terms of design parameters. The main findings from this work are summarized here:
In terms of joint behavior of the variables peak flow-volume:

- The Paraná and Paraguay rivers show that the correlation between the two variables is higher for the events observed during the El Niño phase.
- The Neuquén river shows a higher correlation during La Niña phases, and the difference between the two phases is more important for the summer floods.
- The floods observed in the Limay river do not show any difference for the two phases.

In terms of magnitude of the events:

- Paraná and Paraguay rivers show higher values of peak flow for El Niño phase, in agreement with Camilloni & Barros (2003). The volume series observed in the Paraná river are also higher for the El Niño, whereas for the Paraguay volume series are similar in for both phases.
- The winter floods observed in the Limay and Neuquén show higher values of peak flow and

volume during El Niño phase compared to La Niña. However if the total serie is considered, the floods observed in the Limay river include more critical cases whereas for the Neuquén river the additional floods are more concentrated in low values of peak flow and volume pairs.

- The summer floods observed in the Limay river during La Niña have lower and less variable peak values but higher and more variable volume values. In the Neuquén river if all flood events are considered, more extreme are included in the analysis.

The effect of the joint behavior and magnitude in terms of a design criterion:

- Paraná and Paraguay rivers: El Niño events would result in higher values of peak flow, therefore may be considered for structures in which flow is important for the design criteria.
- Neuquén for winter floods: El Niño events may be used for structures in which volume is important for the design criteria.
- Limay for summer floods: La Niña events may be considered for structures in which peak flow is important for the design criteria.

The aim was to evaluate the presence of changes in the relationship between peak flow and volume resulting from climate variability (in this case ENSO phenomenon). These results help to define the vulnerability of infrastructure to the ENSO phenomenon and then may be used for comparison with simulated extreme flows for climate change scenarios in different catchments in Argentina.

6 ACKNOWLEDGEMENTS

The access to the data used for this work was granted by the Subsecretaría de Recursos Hídricos de la República Argentina. They are gratefully acknowledged.

7 REFERENCES

Asquith, William H. (2011): *Distributional Analysis With L-Moment Statistics Using the R Environment for Statistical Computing*. CreateSpace Independent Publishing Platform, United States.

Camilloni, I. and Barros, V. R. (2003): *Extreme discharge events in the Paraná River and their climate forcing.* Journal of Hydrology 278, pp.94–106.

Chiew, F. H. S. and McMahon, T. A. (2002): *Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability*. Hydrological Sciences Journal 47 (3), pp.505-522.

Chowdhary, H., Escobar, L. A. and Singh, V. P. (2011): *Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data*. Hydrology Research 42.2 (3), pp.193–216.

Delignette-Muller, M. L., Pouillot, R., Denis, J.B. and Dutang, C. (2010): *fitdistrplus: help to fit of a parametric distribution to non-censored or censored data.*

Escalante-Sandoval C. (2007): *Application of bivariate extreme value distribution to flood frequency analysis: a case study of Northwestern Mexico*. Natural Hazards 42, pp.37-46.

Genest, C. and Favre, A. (2007): *Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask.* Journal of Hydrological Engineering 12, pp.347–368.

Hofert, M. and Maechler, M. (2011): *Nested Archimedean Copulas Meet R: The nacopula Package*. Journal of Statistical Software 39(9). (available from http://www.jstatsoft.org/v39/i09/).

Kiem, A. S., Franks, S. W. and Kuczera, G. (2003): *Multi-decadal variability of flood risk.* Geophysical Research Letters 30 (2) 1035.

Kiladis, G. N. and Diaz, H. F. (1989): *Global climatic anomalies associated with extremes in the Southern Oscillation.* J. Climate 2, pp.1069-1090.

Kojadinovic, I. and Yan, J. (2010): *Modeling Multivariate Distributions with Continuous Margins Using the copula R Package*. Journal of Statistical Software 34 (9), pp.1-20. (available from http://www.jstatsoft.org/v34/i09/).

Moyano, C. and R. Diaz (2006): *Mapa argentino de zonas áridas, semiáridas y subhúmedas*. Proyecto del CONAPHI.

Nathan, R. J. and McMahon, T. A. (1990): *Evaluation of Automated Techniques for Baseflow and Recession Analysis.* Water Resources Research 26 (7), pp.1465-1473.

Nelsen, R. B. (2006): *An Introduction to Copulas*. 2nd edition. Springer, New York.

Prieto, M. R. (2007): *ENSO signals in South America: rains and floods in the Paraná River region during colonial times*. Climatic Change 83, pp.39–54.

R Development Core Team (2012): *R: A Programming Environment for Data Analysis and Graphics*. R version 2.15.0 (2012-03-30).

Rakhecha, P. R. and Singh, V. P. (2009): *Applied Hydrometeorology.* Copublished by Springer.

Secretaría de Energía (2004): *Informe del sector eléctrico año 2003 provisorio*. Argentina.

Shiau, J. T. (2003): *Return period of bivariate distributed extreme hydrological events*. Stochastic Environmental Research and Risk Assessment 17, pp.42–57.

Trenberth, K.E. (1997): *The Definition of El Nino*. Bulletin of the American Meteorological Society, 78, pp.2771-2777.

UNEP (1997): *World Altas of desertification*. United Nations Environment Programme. Publisher Middleton N. and Thomas D.

Yan, J. (2007): *Enjoy the Joy of Copulas: With a Package copula.* Journal of Statistical Software 21 (4). (available from http://www.jstatsoft.org/v21/i04/).

Modelling annual maxima of daily rainfall in Madeira Island

Délia Gouveia-Reis[1], Luiz Guerreiro[2] Lopes and Sandra Mendonça[3]

[1]Centre for Exact Sciences and Engineering, University of Madeira, Funchal, Centre of Statistics and Applications, Lisbon, and Mountain Research Centre, Bragança, Portugal, Email: delia@uma.pt

[2]Centre for Exact Sciences and Engineering, University of Madeira, Funchal, Mountain Research Centre, Bragança, and ICAAM/UE, Évora, Portugal

[3]Centre for Exact Sciences and Engineering, University of Madeira, Funchal, and Centre of Statistics and Applications, University of Lisbon, Lisbon, Portugal

Abstract

Madeira Island is located in the Atlantic Ocean off the coast of Northwest Africa, between latitudes 32º30'N–33º30'N and longitudes 16º30'W–17º30'W. Extreme rainfall events have triggered a significant number of flash floods, landslides and debris flows in this volcanic island along its past and recent history. One of the most significant events was the one that happened on the 20th of February 2010, which caused 45 casualties, six missed people and extensive damage to properties and infrastructures. Madeira Island is therefore a natural laboratory for the study of extreme precipitation events and its consequences. In this study, tests based on the likelihood ratio statistic and the probability-weighted moments were used to test the hypothesis of a Gumbel distribution for the annual 1-day maximum rainfall data, from 19 rain gauge stations, provided by the Department of Hydraulics and Energy Technologies of the Madeira Regional Laboratory of Civil Engineering. The rainfall records, with 22, 23 and 31 years of extension, were drawn from rain gauge stations located in the northern and southern hillsides of the island. The estimates for the generalised extreme value distribution (GEV) obtained by the methods of maximum likelihood and probability-weighted moments revealed the influence of the proximity to the sea and altitude on the spatial distribution of extreme rainfall, in addition to the natural differences observed on the windward and lee sides of any mountainous island. Estimates for 50- and 100-year return levels were also obtained. The existence of trends in the parameters' values was also analysed, revealing a significant evidence of a linear trend, both in location and scale parameters, for one location in the north side of the island.

1 INTRODUCTION

Madeira Island is a volcanic island located in the Atlantic Ocean off the Northwest African coast, between latitudes 32º30'N–33º30'N and longitudes 16º30'W–17º30'W, that presents a significant number of rainfall-induced flash floods along its history. There are reports from the 17th century mentioning the occurrence of flash floods (Silva & Menezes, 1945), but the one known to have caused the largest number of casualties, with more than 800 deaths, occurred on the 9th of October 1803 (Fragoso *et al.,* 2012). After that major occurrence, other extreme precipitation events have triggered at least thirty significant flash floods (Quintal, 1999). More recently, the most significant one was the one that happened on the 20th of February 2010, which caused 45 casualties, six missed people and extensive damage to properties and infrastructures, being Funchal and Ribeira Brava the most affected areas (Fragoso *et al.,* 2012; Nguyen *et al.,* 2013). The weather station of Funchal is the oldest in Madeira Island, having started to operate in January 1865 (Silva & Menezes, 1945). Already in 1895, there was the intention of installing another weather station in Pico do Areeiro whose observations, combined with Funchal weather station measurements, would constitute important data for the study of Madeira Island climate and its comparison with other health resort islands (Silva & Menezes, 1945). Its facilities belonged to the General Council of the Autonomous District of Funchal from 1911, but only began providing rainfall and temperature data in November 1936 (Silva & Menezes, 1945). In order to provide useful information for agriculture, more weather stations were settled on the island, at different altitudes, from 1936 to 1955 (Pereira, 1989). Nowadays, Madeira Island is well covered by rain gauge stations maintained by three different

organizations, namely the Portuguese Institute of the Sea and Atmosphere, the Madeira's Investments and Water Management company, and the Regional Laboratory of Civil Engineering (Fragoso *et al.,* 2012). Madeira Civil Engineering Laboratory's Department of Hydraulics and Energy Technologies provided for this study annual 1-day maximum rainfall data from 19 rain gauge stations maintained in the past by the General Council of the Autonomous District of Funchal.

On the other hand, the generalised extreme value distribution (GEV) is widely used for modelling extremes of natural phenomena (cf., e.g., Hosking *et al.,* 1985), and GEV distributions are also used in this work to model the available data. Also in this study, tests based on the likelihood ratio statistic and the probability-weighted moments were used to test the hypothesis of a Gumbel distribution. Two methods, maximum likelihood (ML) and probability-weighted moments (PWM) were used to obtain estimates for the GEV parameters. The 50- and 100-year return level estimates were also obtained and, in addition, we explored the existence of trends in the parameters' values.

2 STATISTICAL ANALYSIS

2.1 Methodology

The generalised extreme value (GEV) family of distributions, that arises as the limiting distribution of the maximum of a series of independent and identically distributed random variables, has the distribution function given by $G_\gamma(\sigma x + \mu)$ where

$$G_\gamma(x) = \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right), & 1+\gamma x > 0, \ \gamma \neq 0 \\ \exp\left(-\exp(-x)\right), & x \in IR, \ \gamma = 0 \end{cases} \quad (1)$$

and γ, μ and σ are, respectively, the shape, location, and scale parameters. This distribution will here be referred to as Model 1 when $\gamma \neq 0$. The particular case of $\gamma = 0$ is the Gumbel distribution and will be referred as Model 2. Since Models 1 and 2 are nested in the GEV family, a model choice can be made applying the likelihood ratio test (Coles, 2001). At the significance level *α*, Model 2 is rejected in favour of Model 1 i2($l_1(M_1)$ - ($l_2(M_2)$) > $\chi^2_{1-\alpha,1}$f where $l_1(M_1)$ and $l_2(M_2)$ are the maximised values of the log-likelihood for Models 1 and 2, respectively. When considering variations of Models 1 and 2 characterised by a linear trend in one or both location and scale parameters, the applied chi-square (1–*α*)-quantile will be $\chi^2_{(1-\alpha,k)}$, where *k* is the number of parameters equal to zero in the sub-model considered. The hypothesis of a Gumbel distribution can also be analysed by a test based on the probability-weighted moments estimate of γ presented by Hosking *et al.* (1985), where the value for the test statistic $\hat{\gamma}$ (n / 0.5633)$^{1/2}$ is compared with the critical values of the standard normal distribution. The good performance of the method of probability-weighted moments for small samples made this method more popular than the maximum likelihood estimation method in applications to hydrologic extremes (Hosking *et al.,* 1985). On the other hand, it is usually more convenient to interpret extreme value models in terms of return levels. The return level estimates, $\hat{q}_p$, are obtained by the estimation of the extreme quantiles of the annual maximum distribution, given by

$$\hat{q}_p = \begin{cases} \mu - \frac{\sigma}{\gamma}\left(1-\left(-\log(1-p)\right)^{-\gamma}\right), & \gamma \neq 0 \\ \mu - \sigma\left(1-\left(-\log(1-p)\right)\right), & \gamma = 0 \end{cases} \quad (2)$$

Where μ, σ, and γ and are replaced by their respective estimates (Coles, 2001).

2.2 Data

The data set analysed in this study, as mentioned before, consists of annual daily maximum rainfall records from 19 rain gauge stations in Madeira Island. Figure 1 shows the location of each station, where the colours (blue, green, orange, and red) of the markers correspond to the four altitude classes considered (<300 m, 300–600 m, 600–900 m, and >900 m, respectively). The markers

identification can be found in Table 1, which also provides additional information about each station used in this study, namely its latitude, longitude and altitude, and the measurement period considered.

Figure 1: Location of the rain gauge stations considered

Table 1: Details of the rain gauge stations

Name (Marker)	Latitude	Longitude	Altitude (m)	Period
Areeiro (A)	32º 43'N	16º 55'W	1610	1950–1980
Bica da Cana (B)	32º 45'N	17º 03'W	1560	1950–1980
Poiso (C)	32º 42'N	16º 53'W	1360	1959–1980
Montado do Pereiro (D)	32º 42'N	16º 53'W	1260	1950–1980
Encumeada (E)	32º 45'N	17º 01'W	900	1959–1980
Ribeiro Frio (F)	32º 43'N	16º 53'W	874	1950–1980
Queimadas (G)	32º 46'N	16º 54'W	860	1950–1980
Porto Moniz (H)	32º 50'N	17º 11'W	653	1950–1972
Ponta do Pargo (I)	32º 47'N	17º 14'W	570	1950–1972
Santo António (J)	32º 40'N	16º 57'W	525	1950–1972
Sanatório (K)	32º 39'N	16º 54'W	380	1950–1980
Santana (L)	32º 48'N	16º 53'W	380	1950–1980
Loural (M)	32º 46'N	17º 02'W	307	1950–1972
Machico (N)	32º 43'N	16º 47'W	160	1959–1980
Ponta Delgada (O)	32º 49'N	16º 59'W	136	1950–1980
Funchal (P)	32º 38'N	16º 53'W	58	1950–1980
Santa Catarina (Q)	32º 41'N	16º 46'W	49	1959–1980
Lugar de Baixo (R)	32º 40'N	17º 05'W	15	1950–1980
Ribeira Brava (S)	32º 40'N	17º 04'W	10	1950–1972

2.3 Results and discussion

The hypothesis of a Gumbel distribution was only rejected with the likelihood ratio test for three locations in the north of the island. The shape parameter estimate is negative for Queimadas (G), a rain gauge located further from the sea than Porto Moniz (H) and Ponta Delgada (O) stations, which present positive shape parameter estimates. All choices resulting from the application of the likelihood ratio test and the corresponding model parameters' estimates, obtained by the application, for each location, of the ismev and extRemes R language packages (R, 2011), are presented in Table 2.

Table 2: Chosen models and maximum likelihood parameter estimates

Station Name (Marker)	Model	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\gamma}$
Areeiro (A)	Model 2	162.49	42.71	-
Bica da Cana (B)	Model 2	128.68	36.86	-
Poiso (C)	Model 2	127.04	34.69	-
Montado do Pereiro (D)	Model 2	129.10	40.00	-
Encumeada (E)	Model 2	158.62	45.89	-
Ribeiro Frio (F)	Model 2	125.33	39.62	-
Queimadas (G)	Model 1	111.70	25.04	–0.209
Porto Moniz (H)	Model 1	57.43	24.41	0.326
Ponta do Pargo (I)	Model 2	60.21	18.60	-
Santo António (J)	Model 2	66.41	27.80	-
Sanatório (K)	Model 2	67.58	24.92	-
Santana (L)	Model 2	89.87	35.75	-
Loural (M)	Model 2	119.59	46.02	-
Machico (N)	Model 2	60.73	27.21	-
Ponta Delgada (O)	Model 1	74.43	28.69	0.236
Funchal (P)	Model 2	47.49	20.29	-
Santa Catarina (Q)	Model 2	52.38	16.89	-
Lugar de Baixo (R)	Model 2	42.16	14.85	-
Ribeira Brava (S)	Model 2	52.38	17.01	-

When testing Model 2 using the test presented by Hosking *et al.* (1985) the results are similar (see Table 3), with the exception of Queimadas (G) and Santo António (J). Table 3 also shows the corresponding model parameter estimates obtained by the use of the *fExtremes* R package (R, 2011).

Table 3: Chosen models and probability-weighted moments parameter estimates

Station Name (Marker)	Model	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\gamma}$
Areeiro (A)	Model 2	96.86	154.02	-
Bica da Cana (B)	Model 2	76.82	123.34	-
Poiso (C)	Model 2	75.33	125.75	-
Montado do Pereiro (D)	Model 2	77.23	128.24	
Encumeada (E)	Model 2	94.68	152.60	-
Ribeiro Frio (F)	Model 2	74.75	127.66	-
Queimadas (G)	Model 2	65.07	97.66	-
Porto Moniz (H)	Model 1	56.51	23.88	0.335
Ponta do Pargo (I)	Model 2	35.83	60.98	-
Santo António (J)	Model 1	70.87	30.37	–0.319
Sanatório (K)	Model 2	40.44	72.15	-
Santana (L)	Model 2	53.89	99.39	-
Loural (M)	Model 2	72.04	124.34	-
Machico (N)	Model 2	36.49	69.44	-
Ponta Delgada (O)	Model 1	73.14	27.50	0.283
Funchal (P)	Model 2	28.49	53.86	-
Santa Catarina (Q)	Model 2	31.30	53.04	-
Lugar de Baixo (R)	Model 2	25.24	43.62	-
Ribeira Brava (S)	Model 2	31.18	54.24	-

We observe that for Ponta Delgada (O) and Porto Moniz (H) data, where we reject Model 2 independently of the method used, the parameter estimates are relatively similar. This is not the case when Model 2 is chosen since the maximum likelihood scale parameter estimates are much smaller than the corresponding probability-weighted moments estimates. The opposite happens with the location parameter estimates, but in a less pronounced way. Although for the distributions corresponding to these cases there is not significant evidence to choose Model 1 in opposition to Model 2, we observe that the corresponding shape parameter estimates are not zero independently of the method applied. Following Coles (2001), who states that the safest option is to accept that there is uncertainty about the value of the shape parameter and to prefer the inference based on the

GEV model whether the Gumbel model is adequate or not, we here chose to deal only with the general GEV distribution. Variations of Model 2 characterised by a linear trend in one or both of the location and scale parameters were also tested. Only in the case of Bica da Cana (B), the rain gauge station located at the highest altitude in the northern side of the island, there is evidence to suggest a linear trend with respect to time at a 0.05 level of significance. More precisely, there is evidence for a linear trend both in location and scale parameters, with $\hat{\mu}(t) = 115.99 + 1.26t$ and $\hat{\sigma}(t) = 2.45 + 0.06t$ On the north side of the island, we observe that the value of the shape parameter for the GEV distribution is positive for the rain gauge stations located nearest to the sea and it is negative for the the stations located in the interior of the island, namely Bica da Cana (B), Encumeada (E) and Loural (M). For the rain gauge stations on the south side, there are cases of positive and negative shape parameter estimates both in the interior and near the coast. Besides the shape parameter estimates, Table 4 presents the maximum likelihood and probability-weighted moment estimates for the location and scale parameters for each station.

Table 4: GEV parameters estimates by ML and PWM

Station Name (Marker)	Method	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\gamma}$
Areeiro (A)	ML	166.07	44.22	–0.153
Areeiro (A)	PWM	165.94	45.69	–0.167
Bica da Cana (B)	ML	131.47	37.07	–0.143
Bica da Cana (B)	PWM	129.40	34.15	–0.033
Poiso (C)	ML	124.81	32.74	0.125
Poiso (C)	PWM	125.15	36.50	0.045
Montado do Pereiro (D)	ML	133.85	42.77	–0.214
Montado do Pereiro (D)	PWM	133.45	45.27	–0.223
Encumeada (E)	ML	167.24	49.10	–0.347
Encumeada (E)	PWM	165.69	51.73	–0.319
Ribeiro Frio (F)	ML	125.98	40.16	–0.031
Ribeiro Frio (F)	PWM	126.04	43.76	–0.071
Queimadas (G)	ML	111.70	25.04	–0.209
Queimadas (G)	PWM	111.99	23.68	–0.214
Porto Moniz (H)	ML	57.43	24.41	0.326
Porto Moniz (H)	PWM	56.51	23.88	0.335
Ponta do Pargo (I)	ML	60.60	18.90	–0.038
Ponta do Pargo (I)	PWM	60.34	20.55	–0.061
Santo António (J)	ML	70.86	29.02	–0.293
Santo António (J)	PWM	70.88	30.37	–0.319
Sanatório (K)	ML	67.96	25.22	–0.028
Sanatório (K)	PWM	67.75	27.04	–0.050
Santana (L)	ML	84.78	36.75	0.021
Santana (L)	PWM	84.73	38.40	0.002
Loural (M)	ML	122.29	46.22	–0.119
Loural (M)	PWM	123.23	46.78	–0.158
Machico (N)	ML	60.49	27.06	0.016
Machico (N)	PWM	59.80	27.62	0.030
Ponta Delgada (O)	ML	74.43	28.69	0.236
Ponta Delgada (O)	PWM	73.14	27.50	0.283
Funchal (P)	ML	46.79	19.77	0.066
Funchal (P)	PWM	46.35	20.32	0.070
Santa Catarina (Q)	ML	53.24	17.39	–0.094
Santa Catarina (Q)	PWM	52.80	17.81	–0.071
Lugar de Baixo (R)	ML	43.33	15.45	–0.145
Lugar de Baixo (R)	PWM	42.68	15.66	–0.091
Ribeira Brava (S)	ML	51.74	16.55	0.070
Ribeira Brava (S)	PWM	51.43	17.28	0.060

Although the altitude appears to be a factor influencing the spatial distribution of extreme rainfall, in general we cannot say that the values of the location parameter estimates increase with altitude. Even though Queimadas (G) rain gauge station presents a higher altitude when compared to Loural

(M), the latter shows a location parameter estimate (of approximately 123 mm) greater than the corresponding value for Queimadas (G). Unlike Loural (M), Queimadas (G) is not located on the E-W oriented orographic barrier in the interior of the island and therefore, besides altitude, the proximity to the sea seems to be a factor influencing the spatial distribution of extreme rainfall. Loural (M) is also near the Encumeada (E) rain gauge station, which presents the highest values for the location and scale parameters estimates. Just below the values observed at Encumeada (E), are the values corresponding to Areeiro (A), the rain gauge station located in the south with the highest altitude in the island. Revealing the natural differences observed on the windward and lee sides of any mountainous island, we have the rainfall data from Sanatório (K) and Santana (L) rain gauge stations. These stations have the same altitude, but Sanatório (K), which is located in the southern part of the island, presents lower values for all the estimated values.

The differences between the northern and the southern part of Madeira Island can also be found in terms of return levels. Illustrating these differences, Figures 2 and 3 present the diagnostic plots for the GEV fit to Ribeira Brava (S) and Ponta Delgada (O) rain gauge data, respectively.

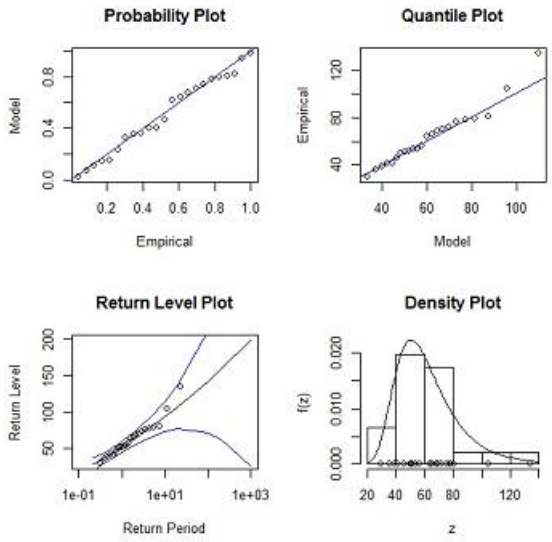


Figure 2: Diagnostic plots for the GEV fit to the Ribeira Brava (S) station data

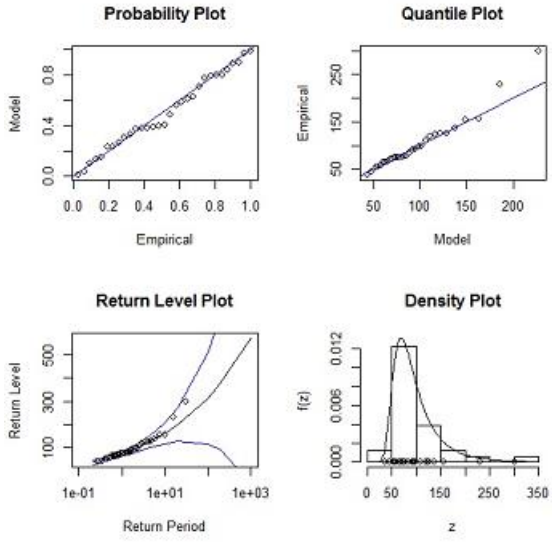


Figure 3: Diagnostic plots for the GEV fit to the Ponta Delgada (O) station data

Table 5 shows the 50- and 100-year return level estimates obtained for each location with the Model 1's parameter estimates produced by ML and PWM methods. For Areeiro (A), Ribeiro Frio (F), Porto

Moniz (H), Ponta do Pargo (I), Santo António (J) and Ribeira Brava (S), the 50- and 100-year return values calculated by both methods are approximately similar ($\hat{q}_{0.02}^{ML} \approx \hat{q}_{0.02}^{PWM}$ and $\hat{q}_{0.01}^{ML} \approx \hat{q}_{0.01}^{PWM}$, respectively). The largest parameter estimates for Poiso (C), Queimadas (G), Santana (L) and Loural (M) were obtained by the maximum likelihood method. For all the rest of rain gauge stations, the higher values for the return levels were found when calculated by the method of probability-weighted moments.

Table 5: Estimates for 50- and 100-year return levels

Station Name (Marker)	$\hat{q}_{0.02}^{ML}$	$\hat{q}_{0.02}^{PWM}$	$\hat{q}_{0.01}^{ML}$	$\hat{q}_{0.01}^{PWM}$
Areeiro (A)	296.09	297.05	312.24	312.77
Bica da Cana (B)	242.23	254.33	256.29	275.02
Poiso (C)	289.29	280.82	328.12	311.66
Montado do Pereiro (D)	247.01	251.50	259.05	263.78
Encumeada (E)	272.22	281.18	280.09	290.52
Ribeiro Frio (F)	273.50	275.31	298.05	297.95
Queimadas (G)	178.53	174.60	185.73	181.26
Porto Moniz (H)	249.67	249.71	317.92	318.01
Ponta do Pargo (I)	129.07	129.07	140.28	140.28
Santo António (J)	138.28	138.66	144.11	144.14
Sanatório (K)	161.09	163.62	176.69	178.89
Santana (L)	249.39	239.26	284.97	266.50
Loural (M)	266.48	259.36	285.93	276.02
Machico (N)	169.47	174.08	189.70	195.98
Ponta Delgada (O)	258.24	268.99	312.96	332.94
Funchal (P)	134.70	137.48	152.94	156.56
Santa Catarina (Q)	110.05	113.56	118.20	122.77
Lugar de Baixo (R)	89.38	94.12	95.21	101.55
Ribeira Brava (S)	125.95	125.93	141.50	141.48

In the northern part of Madeira Island, all return level estimates calculated are higher than 240 mm, excluding the values corresponding to Queimadas (G). In turn, on the south side of the island the return values are less than 180 mm for all rain gauge stations with altitudes below 600 m, with the exception of Machico (N). For Areeiro (A), Poiso (C), Montado do Pereiro (D) and Ribeiro Frio (F), the 50- and 100-year return level estimates are greater than 245 mm.

The proximity between the return value estimates for Queimadas (G) and Machico (N) rain gauge stations suggests the proximity to the sea as a factor to be taken into account in the study of return levels, in addition to natural factors such as altitude or location in the northern or southern part of the island. These two rain gauge stations have distinct altitudes and are located at different but nearby hillsides. In the southwest, we observed another pair of rain gauge stations, Ponta do Pargo (I) and Ribeira Brava (S), with similar return values but closer parameter estimates. Although the distance between rain gauge stations might seem an influential factor, closer rain gauge stations does not mean similar return level estimates. To exemplify this, we may mention the set of three rain gauge stations located in Funchal municipality, namely Santo António (J), Sanatório (K) and Funchal (P).

3 CONCLUSIONS

In this work, an application of extreme value theory to the annual 1-day maximum rainfall data from 19 rain gauge stations in Madeira Island was presented. Although most of the rain gauge stations considered are deactivated, and consequently the rainfall time series are relatively short, it is important to analyse all the available data, given the number of major flash flood events reported so far in Madeira Island. The most recent of these significant events occurred on the 20th of February 2010, with 146.9 mm observed in Funchal and 333.8 mm in Areeiro (Fragoso *et al.,* 2012). Given the GEV estimates obtained in this work by ML, these values correspond to return periods of approximately 79 and 292 years, respectively, or 70 and 297 years when GEV estimates by PWM are used. Although for almost all rainfall data series there was not significant evidence to choose the GEV distribution in opposition to the Gumbel distribution, GEV parameters estimates by ML and

PWM methods were provided in this work given that the corresponding shape parameters estimates are not zero and that location and scale GEV estimates are relatively similar for both methods. The hypothesis of a Gumbel distribution was tested by the likelihood ratio test and by the test presented by Hosking *et al.* (1985), and the same conclusions were obtained for both methods, with the exception of the data from Queimadas (G) and Santo António (J) rain gauge stations. A significant evidence for a linear trend in location and scale parameters was found in the data from Bica da Cana (B), the rain gauge station located at the highest altitude in the northern side of the island. Estimates for the 50- and 100-year return levels were also determined for all rain gauge stations data using the GEV parameter estimates obtained from both methods.

The parameter and return level estimates, regardless of the method used to obtain them, suggest a complex characterization of the spatial distribution of extreme rainfall in Madeira Island. It seems that there is a simultaneous influence of factors such as altitude, proximity to the sea, distance, and location in nearby hillsides or in the northern or the southern part of the island. Besides that, there are differences in the return levels estimates according to the method applied, except for Areeiro (A), Ribeiro Frio (F), Porto Moniz (H), Ponta do Pargo (I), Santo António (J) and Ribeira Brava (S) rain gauge stations. Nevertheless, it can be observed that the 50- and 100-year return levels estimates are greater than 245 and 255 mm, respectively, for all the seven rain gauge stations located farther from the sea, namely Areeiro (A), Bica da Cana (B), Poiso (C), Montado do Pereiro (D), Encumeada (E), Ribeiro Frio (F) and Loural (M). The same can be observed for three more stations, Porto Moniz (H), Santana (L) and Ponta Delgada (O), that are closer to the sea but located in the northern part of the island. There is proximity between the estimates' values for Queimadas (G) and Machico (N) data, although Queimadas (G) and Machico (N) rain gauge stations belong, respectively, to the northern and southern parts of the island. These two stations are located in nearby hillsides, as Ponta do Pargo (I) and Ribeira Brava (S) that are located in the southwest and also present similar return value estimates. For all the rest of the rain gauge stations located in the southern part of the island the 50- and 100-year return levels estimates are smaller than 164 mm and 179 mm, respectively. The rain gauge stations that present the highest and the smallest return level estimates, Areeiro (A) and Lugar de Baixo (R), are both located in the south side of Madeira Island.

4 ACKNOWLEDGMENTS

This work was partially supported by the Portuguese Foundation for Science and Technology (FCT) through the PhD grant SFRH/BD/39226/2007, and by project PEst-OE/MAT/UI0006/2011. The authors wish to thank the Department of Hydraulics and Energy Technologies of the Madeira Regional Laboratory of Civil Engineering, and namely to Dr. Carlos Magro, for providing the rainfall data used in this study, and to the University of Madeira for the logistic support.

5 REFERENCES

Coles, S.G. (2001): *An Introduction to Statistical Modeling of Extreme Values*, Springer, London.

Fragoso, M., Trigo, R.M., Pinto, J.G., Lopes, S., Lopes, A., Ulbrich, S. and Magro, C. (2012): *The 20 February 2010 Madeira flash-floods: Synoptic analysis and extreme rainfall assessment*. Nat. Hazards Earth Syst. Sci., 12, 715-730.

Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985): *Estimation of the generalized extreme-value distribution by the method of probability-weighted moments*, Technometrics, 27, 251–261.

Nguyen, H.T., Wiatr, T., Fernández-Steeger, T.M., Reicherter, K., Rodrigues, D.M.M. and Azzam, R. (2013): *Landslide hazard and cascading effects following the extreme rainfall event on Madeira Island (February 2010)*, Nat. Hazards, 65, 635–652.

Pereira, E.C.N. (1989): *Ilhas de Zarco, Câmara Municipal do Funchal*, Funchal, Portugal.

Quintal, R. (1999): *Aluviões da Madeira; séculos XIX e XX*, Territorium, 6, 31–48.

R Development Core Team (2011): *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Silva, F.A.S. and Menezes, C.A. (1945): *Elucidário Madeirense*, Vols. I and III, DRAC, Funchal. Funchal.

Seasonal Extreme Value Statistics for Precipitation in Germany

Madlen Fischer[1], Henning W. Rust[1] and Uwe Ulbrich[1]

[1]Institute of Meteorology, FU Berlin, Germany, Email: madlen.fischer@met.fu-berlin.de

Abstract

Extreme precipitation events have a strong influence on environment, society and economy. They can lead to floods, mudslides, increased erosion or hail which in turn can cause significant damage. Standard annual return levels are commonly used for the design of hydraulic structures or for risk assessments of insurance companies. Seasonally or monthly resolved return levels have not yet entered the common practice, although they do provide additional information in multiple respect: a higher temporal resolution can be useful for risk management, e.g., for the agricultural or tourist sector, and, due to the larger amount of available data, derived annual return levels are potentially more accurate. We calculate monthly resolved return levels for 1208 stations in Germany using monthly maxima of daily precipitation amounts. The monthly block size is sufficient for a description of monthly maxima with a non-stationary generalized Extreme Value Distribution (GEV). The temporal variation of the location and scale parameter is modeled with harmonic functions and the shape parameter is held constant in time. In this way, parameter estimation benefits from the fact that return levels vary smoothly throughout the year. This approach is more natural and robust than describing each month's maxima separately with a GEV. We use cross validation to determine the order of harmonic functions for all stations. The cross validation error of the favourite model has been reduced by 33 % with respect to the baseline approach based on the individual months.

1 INTRODUCTION

The recent flooding events in Central Europe in May/June 2013 illustrate the strong influence of extreme events like heavy precipitation on environment, society and economy. Furthermore, the frequency of occurrence of extreme precipitation shows trends in several regions of the world due to the global warming (Trenberth *et al.*, 2007; Meehl *et al.*, 2007). Annual maxima statistics are commonly used to describe the return levels which are the basis for, e.g., the design of hydraulic structures or for risk assessments of insurance companies. To characterize the seasonality of heavy rainfall events, we use monthly resolved maxima statistics. On the one hand, this provides additional information for risk assessment, e.g., for the agricultural or tourist sector. On the other hand, derived annual return levels are potentially more accurate due to the smooth variation of return levels throughout the year and the larger amount of available data.

A frequently used concept to analyze the occurrence and magnitude of extremes is extreme value statistic (EVS) (Coles, 2001; Beirlant *et al.*, 2004) with a widespread use in hydrology and climatology (Katz *et al.*, 2002). With a non-stationary EVS approach, the seasonality of extreme precipitation events in the UK has been investigated based on monthly maxima statistics in several studies (Rust *et al.*, 2009; Maraun *et al.*, 2009a; Maraun *et al.*, 2009b).

Section 2 gives an overview over the data used and describes the methods including the calculation of return levels with the block maxima approach of the EVS. It furthermore briefly explains the cross validation used for selecting an adequate seasonal model. In section 3, the return levels for the example station Benediktbeuern and the seasonality of the 100-year return level for all stations are discussed. A short simulation study to demonstrate the increased accuracy (bias and variance) of annual return levels due to using monthly resolved data is outlined in section 4. Section 5 summarizes and gives a conclusion.

2 DATA AND METHODS

2.1 Data basis

The basis for our analysis are daily precipitation amounts measured with Hellmann rain gauges (accuracy of 0.1 mm) for almost 5700 stations from the National Climate Data Centre of the German Weather Service (Deutscher Wetterdienst, https://werdis.dwd.de). Here, we consider a subset of 1208 records covering the period from January 1951 to December 2010 (missing values within this range are allowed). For months with less than five days of rainfall, a maximum is not obtained and the record is considered as missing value. The station locations are shown in Fig. 1. A subset of long records from January 1900 to December 2010 (Fig. 1, black triangles) is used for model selection. The station Benediktbeuern (also a long record) will be discussed in detail (Fig. 1, violet triangle).

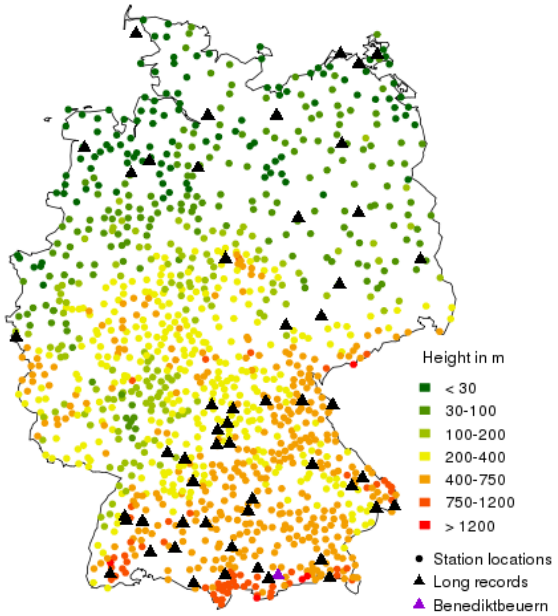


Figure 1: Station locations with records starting at least January, 1st, 1951 and ending earliest December, 31, 2010. Among those, a subset of 54 stations with long records (from January, 1st, 1900 to December, 31st, 2010) used for model selection (black triangles). The example station Benediktbeuern is depicted as violet triangle

2.2 Modeling seasonal extreme precipitation

EVS provides several concepts to describe extreme values. The two main approaches are: 1) the peak-over-threshold approach (POT), which consider excesses over a given threshold and uses the generalized Pareto distribution to model these excesses, and 2) the block maxima approach, which considers maxima out of a certain time block and uses the generalized extreme value distribution (GEV) to describe these. Here, we use block maxima and allow the GEV parameter to vary smoothly in time. This variation is modeled by a series of harmonic functions (sine, cosine) of different orders.

Basis for the approach are maxima *M* taken from a series of independent and identically distributed random variables *X,* e.g., a sequence of daily precipitation amounts with a sufficient length or block size *n.*

$$M_n=\max\{X_1,\ldots, X_n\} \quad (1)$$

In the present study, we use monthly maxima of daily precipitation and verify the adequacy of the block size by quantile plots (not shown). According to the Fisher-Tippet theorem (e.g, Coles, 2001) either the Gumbel, Weibull or Fréchet distribution is suitable to describe the probability distribution of the maxima M_n if a sufficient block size *n* is reached. These three distributions can be summarized in the generalized extreme value distribution GEV

$$G(z)= \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)^{-1/\xi}\right]\right\} \quad (2)$$

defined on $\left\{z : 1+\frac{\xi(z-\mu)}{\sigma}>0\right\}$, where $-\infty< \mu< \infty$, $\sigma > 0$ and $-\infty< \xi<\infty$. The location parameter μ determines the position of the probability distribution function (PDF) and the scale parameter σ specifies its width. The type of distribution is determined by the shape parameter. For the case of a positive shape parameter, $\xi > 0$, the PDF decays slowly (i.e. algebraic) towards large values (Fréchet). For a negative shape parameter, $\xi < 0$, the GEV has a finite upper bound (Weibull). The limiting case $\xi = 0$ describes the light tailed Gumbel distribution with a fast (i.e. exponential) decay towards large values (e.g., Coles, 2001; Embrechts, 1997).
Due to the seasonality of extreme rainfall maxima, a suitable distribution should vary in time. We thus use a non-stationary GEV with time-dependent parameters. A series of harmonic functions provide a natural way to specify seasonality. The location parameter can thus be expressed as

$$\mu_t= \mu_0+ \sum_{n=1}^{N} \mu_1 \cdot \sin(n\omega c_t) + \mu_2 \cdot \cos(n\omega c_t) \quad (3)$$

with t=1,…,12 denoting the month in the year, c_t the centre of the t-th month given in number of days starting from January, 1st, ω=(2π)/365.25 the angular frequency of earth´s and N the amount of orders (Maraun *et al.*, 2009; Rust *et al.*, 2009; Maraun *et al.*, 2009a). Analogously, the seasonal variations of the scale parameter can be modeled with

$$\sigma_t= \sigma_0+ \sum_{n=1}^{N} \sigma_1 \cdot \sin(n\omega c_t) + \sigma_2 \cdot \cos(n\omega c_t) \quad (4)$$

In principle, the shape parameter could be described with a seasonal cycle as well, but because it is difficult to estimate already in the stationary case and the interference is correlated with the scale parameter (Ribereau *et al.*, 2011), we decided to leave it constant in time

$$\xi_t=\xi_0 \quad (5)$$

The parameters of the GEV are estimated using maximum-likelihood (e.g., Coles, 2001). Aim of an extreme value analysis is an estimate of return levels r_T for given return periods *T, i.e.* the level (precipitation amount) which is expected to be exceeded on average once in the return period *T*. The return level is thus a quantile of the GEV for the probability *p=1-1/T*.

$$r_T= \begin{cases} \mu-\frac{\sigma}{\xi}\left[1-y_T^{-\xi}\right] & \text{for } \xi\neq 0 \\ \mu-\sigma \log y_T & \text{for } \xi\rightarrow 0 \end{cases} \quad (6)$$

with $y_T= -\log(1-1/T)$. The return period $T=\frac{1}{1-p}$ is related to the probability 1-*p* of a maximum exceeding the return level r_T. In an engineering context, the 100-yr return level, which is expected to be exceeded on average once every 100 years, is frequently used for dimensioning of certain structures, such as bridges, dams or urban drainage system. The confidence intervals for the return levels can be derived from the GEV parameters using the Delta method (Coles, 2001).

2.3 Model selection

We use cross validation (CV, Wilks, 2011) to guide the selection of an adequate order of harmonic functions for μ and σ. The main idea of CV is to split the data into two or more parts to estimate a prediction error of the model. In a first step the model parameters are estimated on one part of the data and the prediction error is obtained on the (independent) other part. The process is repeated until a prediction error has been obtained for every data point. Here, we split the data of 54 stations with long records (Fig. 1, black triangles) into three random parts including 37 years. For the location and scale parameter we consider harmonic functions up to order five. Overall the model selection covers 36 variants of a non-stationary model. As a measure for the cross validation error (CVE), we use the negative log-likelihood. In addition, we analyze a baseline model, characterizing the maxima of each month individually, ending up with 12 stationary GEVs, one for each month. Figure 2 shows

the mean CVE over all 54 records for all 36 model variants. The first position of the model name on the x-axes represent the order of the location parameter μ and the second the order of scale parameter σ. The best model is marked in green and the baseline model BM with a red triangle.

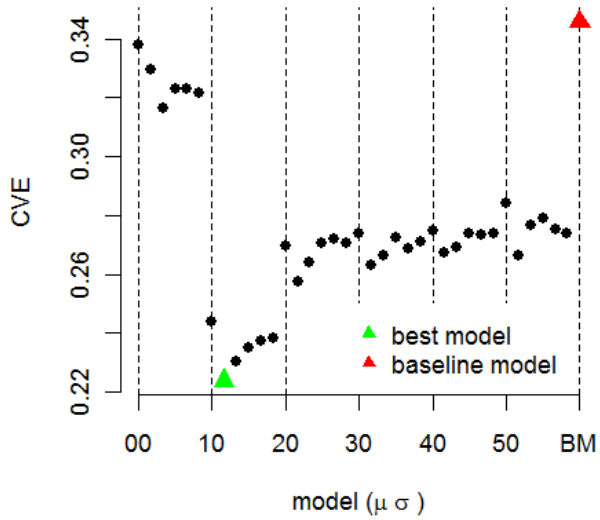


Figure 2: Mean CVE over the 54 long records use for model selection for 36 non-stationary models (orders of 0 up to 5 for location and scale parameter) and the mean CVE from the baseline model BM (red triangle) characterizing each individual month. The best model is marked with a green triangle

The CVE of the baseline model (red triangle) is larger than the value of the worst non-stationary model. The best value is reached for a seasonal model with first order in location and scale (green triangle). As the second best model is not much worse and it offers more flexibility needed for particular stations, we decided to take the second order of the scale parameter also into account. For the sake of having just one model, we take into account that for some stations the model is over-parameterized. However, the CVE of the chosen model has been reduced by 33 % with respect to the baseline model. We thus end up with a seasonal model with a first order harmonic function in location (3 parameters to estimate), a second order in scale (5 parameters to estimate) and a time constant shape parameter (1 parameter to estimate) This model is abbreviated as 120-model. The parameters can be expressed as follows:

$$\mu_t = \mu_0 + \mu_1 \cdot \sin(\omega c_t) + \mu_2 \cdot \cos(\omega c_t) \quad (7.1)$$

$$\sigma_t = \sigma_0 + \sigma_1 \cdot \sin(\omega c_t) + \sigma_2 \cdot \cos(\omega c_t) + \sigma_3 \cdot \sin(2\omega c_t) + \sigma_4 \cdot \cos(2\omega c_t) \quad (7.2)$$

$$\xi_t = \xi_0 \quad (7.3)$$

For this seasonal model 9 parameters have to be estimated. This should be confronted with the baseline model with 36 parameters (3 parameters for one month).

3 SEASONAL RETURN LEVELS

3.1 Example station Benediktbeuern

To illustrate the seasonality of heavy rainfall the return levels corresponding to several return periods (probabilities) are exemplarily presented for the station Benediktbeuern in the foothills of the Alps (Fig. 1, violet triangle). The record covers a period of nearly 120 years, from 01.09.1891 to 31.12.2011. Figure 3 shows the monthly maxima of observed daily precipitation amounts as Box-Whisker-Plots for each month (light grey). The black line within the box depicts the median, the upper and lower end of the box the 0.75 and 0.25 quantile, respectively. The whiskers describe the

maximum and minimum unless they are outside the 1.5 interquartile range. The data points outside of the error bars are depicted explicitly. There is strong evidence for significance of differences in medians, if the notches do not overlap. In addition, the 0.99 quantile (corresponding to a 100-year return period) of the observed data is shown as a black line above the Box-Whisker-Plots. The seasonal cycle of extreme precipitation in Benediktbeuern is characterized by a maximum in summer and a minimum in winter. At other stations, the behavior might be quite different.

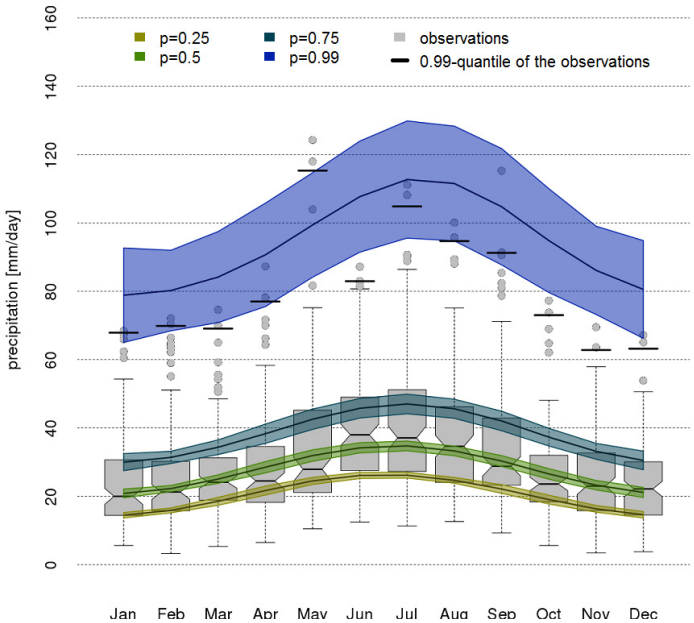


Figure 3: Box-Whisker-Plots (light grey) of the monthly maxima of observed daily precipitation amounts for Benediktbeuern (01.09.1891 - 31.12.2011) and the return levels derived from the seasonal model (colored lines) for p= 0.25, 0.5, 0.75, 0.99 (from bottom to top, yellow, green, turquoise, blue) with their 95%- confidence intervals (shaded area). The black lines above the error bars denote a model-free 0.99 quantile obtained just by ranking the observed data

In addition, the return levels corresponding to *p*=0.25 (yellow), 0.5 (green), 0.75 (turquoise) and 0.99 (blue) derived from the seasonal model are depicted in Fig. 3 by colored lines with corresponding 95% confidence intervals (shaded areas) as well. The 0.25-, 0.75- and 0.5 quantiles well describe the lower and upper quartile and the median of the observed daily precipitation amounts, respectively. A rough guide for the quality of the 100-year return levels is to count the precipitation events above the 0.99 quantile line. Because of an observation period of nearly 120 years in Benediktbeuern, 1.2 exceedings per month or rather 14.2 exceedings per year are expected. As shown in Fig. 3 only 4 daily precipitation events exceed the 0.99 quantile, but with the involvement of the lower confidence interval the model is up to describe the once-in-a-hundred-years events quite well. On the other hand the 0.99 quantile determined by ranking the observed data (black lines) underestimate the 100-year return levels.

3.2 100-year return level for all considered stations

This section covers the seasonal and spatial variations of the 100-year return levels for all 1208 stations. Figure 4 depicts the return levels in mm/day for each month. Because of increased elevation of air masses and accumulation effects at the hillside, extreme precipitation events are more intense in higher altitudes. In January the highest precipitation events reach values from 97,5 mm/day to 116,2 mm/day in the Alps and the Black Forest. The return levels increase from south to north with chronological sequence until the maximum values of more than 116,3 mm/day are reached in summer in the Alps. In general, the differences between the precipitation amounts in lower and higher altitudes are not as pronounced as in winter. Up to December the return levels decreased from north to south and offer nearly the same pattern as in January. Most of the stations show seasonality with less intense extremes in winter and more intense extremes in summer, but in some regions, for example the Black Forest or the Harz, the seasonal cycle indicates the opposite so that the patterns in summer are more similar.

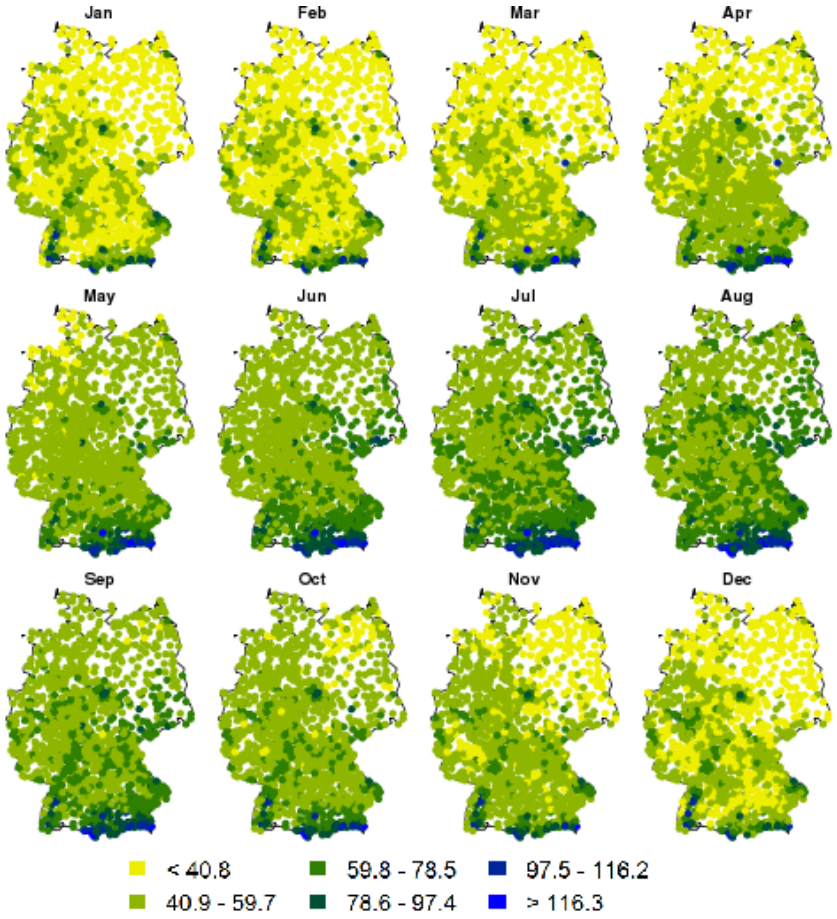


Figure 4 100-year return levels in mm/day conditioned on the month of their occurrence for 1208 stations (dots). The panels show the months January to December left to right and top to bottom

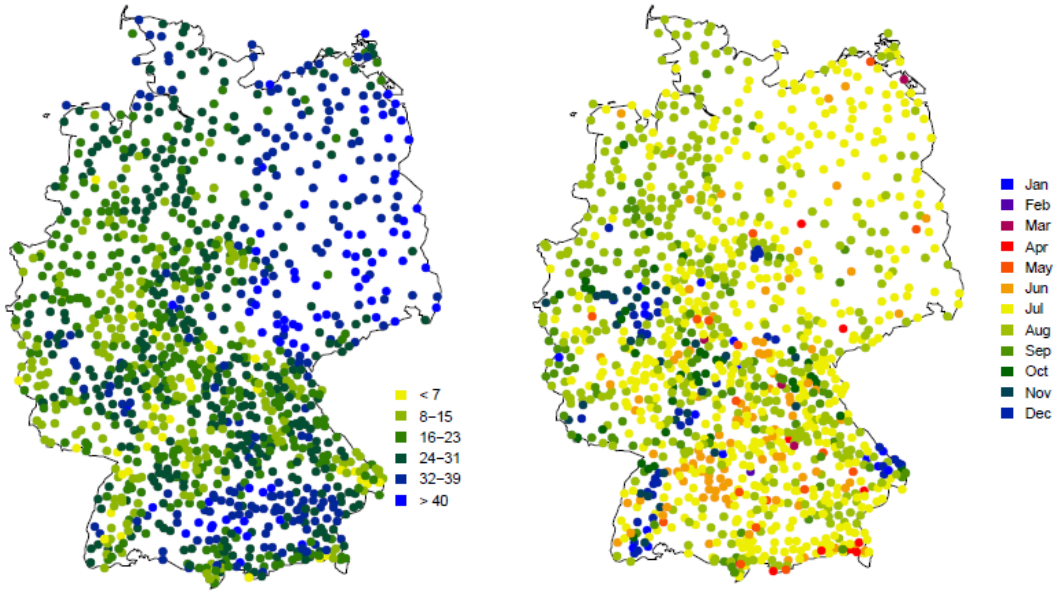
In addition to the return levels for each month, the amplitude and phase of the seasonal variations of precipitation is an interesting quantity. As mean return levels vary in space, it is necessary to relate an amplitude to a characteristic quantity, such as a mean return level. The amplitude can thus be expressed as:

$$A_{\%} = \frac{RR_{max} - RR_{min}}{RR_{max} + RR_{min}} \cdot 100\% \tag{8}$$

with RR_{max} denoting the maximum and RR_{min} denoting th minimum.

In Fig. 5 a) the amplitudes are depicted for all stations. Small/large values indicate weak/strong seasonality. The fairly high values at the stations in the east and in the valleys point out that the intra-annual variations of extreme precipitation are very distinctive. These regions are likely to be influenced by heavy convective precipitation in summer, whereas the effect of all-the-year stratiforme precipitation is more dominant at stations further west with a more maritime influence. The weakest seasonality can be found at stations at high altitudes, where a lift of air masses and accumulation effects lead to all-the-year intensive extreme precipitation. Finally, the month of occurrence of the maximum 100-return level is depicted in Figure 5b). In particular for the agricultural and tourist sector the most endangered regions together with the months of most intense extremes are of special interest. In most areas the most intense precipitation are found in summer. At many stations further west, the highest return levels occur from September to November. This could be an indicator for the influence of increased cyclone activities in autumn. Some areas with higher altitudes, i.e. the Harz, the Black Forest and the Bavarian Forest, are characterized by a precipitation maximum in the winter months. In addition to station altitude the main wind direction is an important characteristic. If a mountain range is perpendicular to the main wind direction (west, strongest storminess in winter) the lifting effects and the connected convection are maximal and lead

to high precipitation amounts especially in winter.



a) b)

Figure 5: a) Amplitude of intra-annual variations of the 100-year return level. Large/small amplitudes indicate a strong/weak seasonality. b) Month of the occurrence of the maximum 100-year return level for 1208 stations

4 ANNUAL RETURN LEVELS

For many applications in the field of hydraulic design and risk assessment, the annual return levels are the values of interest. Commonly the stationary GEV (Eq. 2) is used to describe annual maxima and to estimate annual return levels. However, annual return levels can also be determined based on the seasonal model. The *T*-year annual return level r_T can be derived from the non-stationary model by numerically solving the following equation:

$$\prod_{i=1}^{12} G_i(r_T) = 1 - \frac{1}{T} \tag{9}$$

with $G_i(r_T)$ being the probability of the occurrence of a value smaller than r_T in the month i. To illustrate the improvement of obtaining annual return levels derived from the seasonal model compared to directly modeling annual maxima with the GEV, we carry out the following simulation study: We generate series of monthly maxima using GEV-distributed random variables with a typical seasonal component in location and scale. On the basis of these monthly maxima series we apply two strategies for deriving annual return levels: 1) obtain annual maxima and estimate parameters for a stationary GEV, and 2) model the annual maxima with a seasonal GEV model. We generate series of different length (12 years to 600 years), 10000 iterations each. The results of both strategies are compared to the model used for generation. Fig. 6a) shows the variance and bias for the annual return levels obtained using 1) (black) and 2) (red) for different return periods. Both, variance and bias of annual return levels derived from the seasonal model are significantly smaller than for the standard approach based on annual maxima., particularly for the important cases with short record lengths and long return periods. The improvement of the annual return levels derived from the seasonal model is reflected by much smaller confidence intervals. Annual 100-year return levels determined with the 120-model are depicted in Figure 6b).

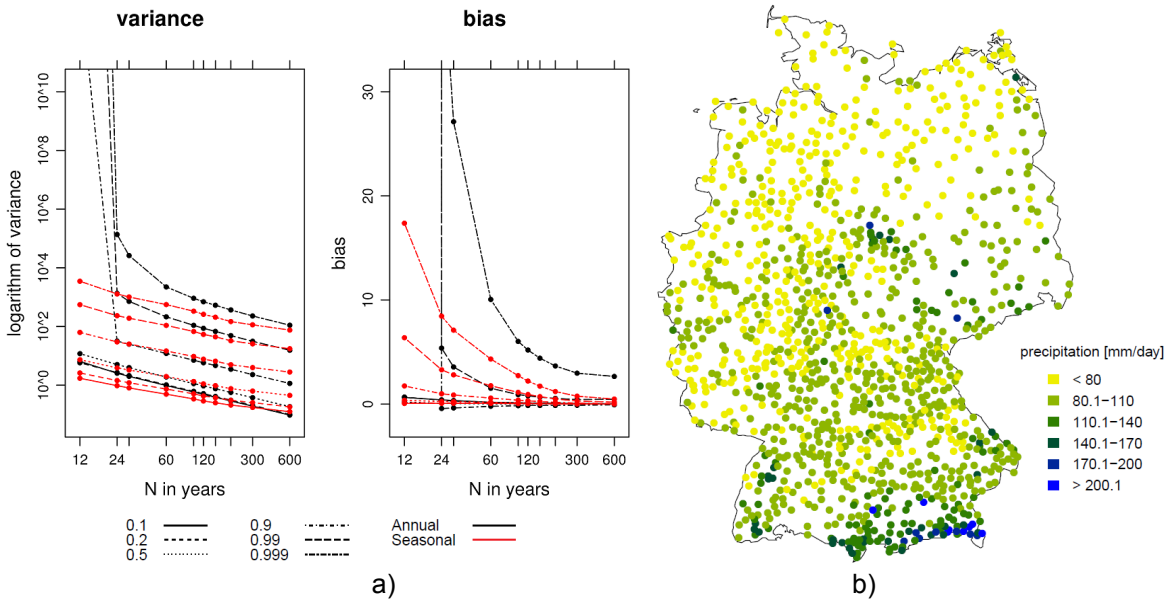


Figure 6: a) Variance and bias for annual return levels derived on the basis of annual maxima (black) and derived from monthly maxima using the season model (red) for several different return periods. On the x-axis are different length of observation (12 to 600 years). b) Annual 100-year return level derived from the seasonal 120-model for 1208 stations (dots)

5 SUMMARY AND CONCLUSION

We show that block maxima modeling can be applied to monthly maxima of daily precipitation amounts and seasonality can be explicitly resolved. Seasonality is modeled using harmonic function of different orders for the location and scale parameter of the GEV, the shape parameter is held constant in time. For selecting adequate orders for the harmonic functions, we use cross validation with the negative log-Likelihood measuring forecast error. First order harmonic functions in location and scale yield the lowest CVE. However, as we aim for one model suitable for all stations, we decided to include also the second order harmonics for the scale parameter, a model with only a slightly larger CVE but more flexibility which is required for a significant number of stations. The cross validation error of the chosen model (120-model, first order in location, second order in scale and constant shape) is about 33 % smaller than the CVE for the baseline model. Overall the seasonal model has just 9 parameters (3 for location, 5 for scale and 1 for shape) whereas the baseline model has 36 parameters.

Additionally to resolving seasonality, we address the characteristics of two ways of estimating annual return levels: 1) based on annual maxima described with the GEV and 2) based on monthly maxima using the seasonal GEV model. To this end, we carry out a simulation study, which demonstrates the improvement in annual return levels derived from the seasonal model with respect to the return levels based on annual maxima. Variance and bias of the two approaches are compared and it turns out that the annual return levels derived from the seasonal model and monthly maxima are more accurate especially for short observation periods and long return periods.

This is plausible as there is additional information available, such as monthly instead of annual maxima and the idea of smoothly varying return levels throughout the year. It is thus consequent, that annual return levels can be estimated more accurately with the approach presented here.

6 ACKNOWLEDGMENTS

We thank the German Weather Service for providing the observed precipitation records.

7 REFERENCES

Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J. De Waal, S. and Ferro, Ch. (2004): *Statistics of Extreme – Theory and Applications*, John Wiley & Sons, Ltd.

Coles, S. (2001): *An Introduction to Statistical Modeling of Extreme Values.*, Springer-Verlag London.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997): *Modelling Extremal Events for Insurance and Finance. Applications in Mathematics*. Springer, Berlin.

Maraun, D., Rust, H.W. and Osborn, T.J. (2009): *Synoptic airflow and UK daily precipitation extremes – Development and validation of a vector generalised linear model.*, Clim Dym 36, pp.261-275

Maraun, D., Rust, H.W. and Osborn, T.J. (2009): *The annual cycle of heavy precipitation across the United Kingdom: a model based on extreme value statistics.*, Int. J. Climatol 29, pp.1731-1744.

Meehl, G.A., Stocker, TF., Collins, W.D., Friedlingstein, P., Gaye, A.T., Gregory, J.M, Kitoh, A., Knutti, R., Murphy, J.M., Noda, A., Raper, S.C.B., Watterson, I.G., Weaver, A.J. and Zhao, Z.C. (2007): *Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Chapter Global Climate Projections*. Cambridge University Press: Cambridge, New York.

Ribereau, P., Naveau, P., and Guillou, A. (2011): *A note of caution when interpreting parameters of the distribution of excesses*, Adv. Water Resour. 34, pp.1215.1221.

Rust, H.W., Maraun, D. and Osborn, T.J. (2009): *Modelling seasonality in extreme precipitation - A UK study*, Eur. Phys. J. Special Topics 174, pp.99-111.

Trenberth, K.E., Jones, P.D., Ambenje, P., Bojariu, R., Easterling, D., Tank, A.K., Parker, D., Rahimzadeh, F., Renwick, J.A., Rusticucci, M., Soden, B. and Zhai, P. (2007): *Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Chapter Observations: Surface and Atmospheric Climate Change.* Cambridge University Press: Cambridge, New York.

Wilks, D.S. (2011): *Statistical Methods in the Atmospheric Science*, Elsevier Inc., third edition.

Extreme precipitation in a changing climate: A regional POT approach

M. Roth[1] and T. A. Buishand[1]

[1]Royal Netherlands Meteorological Institute, De Bilt, Netherlands, Email: m.roth@tue.nl

Abstract

Projections of extreme precipitation are of great importance, considering the potential severe impacts on society. To make inference about future change in extreme precipitation in the Netherlands and north-western Germany, a recently developed regional, non-stationary peaks-over-threshold approach is applied to daily precipitation from two transient simulations of the RACMO2 regional climate model for the period 1950-2100. A temporally varying threshold is used to account for changes in the frequency of precipitation extremes. The height of the threshold is determined by averaging threshold stability plots regionally. The marginal distributions of the excesses are described by generalized Pareto distributions. The parameters of these distributions may vary over time and their spatial variation is modeled by the index flood approach. This non-stationary approach results in smaller confidence bands for the return levels than a moving window approach. Bias-corrected projections of the 50-year return level are presented.

1 INTRODUCTION

Consensus is growing, that the characteristics of extreme precipitation may alter owing to climate change. In order to project the change in extreme precipitation, climate model data have been analyzed and compared to observations. Extreme-value distributions have been fitted to the extremes for two subsets of the data representing current (e.g. 1980-2010) and future (e.g. 2070-2100) climate, assuming stationarity within the time slices and the differences between the two periods have been studied, see e.g. Fowler *et al.* (2005). However, considering only two time slices does not give a picture of the evolution of the extremes, which is e.g. necessary if one is interested in the risk of failure of a hydraulic structure during its expected lifetime. Moreover, the selection of the time slices introduces additional uncertainty. A small shift of the time slices may have large influence on the estimated change. As an alternative, extreme value distributions with time-dependent parameters, which allow the consideration of the full time period, have been used, see e.g. Coles (2001), El Adlouni *et al.* (2007), and Kyselý *et al.* (2010).

The estimation of changes in rare extremes is subject to large uncertainty. A general way to reduce the estimation uncertainty is regional frequency analysis (RFA), where the similarities between different sites in a region are exploited (Hosking & Wallis, 1997). RFA is mostly applied to (annual) block maxima (BM). An alternative to BM is to consider all peaks over a (high) threshold (POT), which is often preferable, owing to the more efficient use of the data.

A regional peaks-over-threshold model, combining the RFA approach and POT data, which can be used to analyze precipitation extremes in a changing climate, was developed by Roth *et al.* (2012). In this model a temporally varying threshold, which is determined by quantile regression, is used to account for changes in the frequency of precipitation extremes. The marginal distributions of the excesses are described by generalized Pareto distributions (GPD), with parameters, that may vary over time and their spatial variation is modeled by the index flood (IF) approach. The core of the index flood (IF) assumption is that the data have a common distribution, after scaling by a site dependent index variable (or index rainfall), see e.g. Hosking & Wallis (1997).

Daily precipitation for the period 1950-2100 from two simulations of the RACMO2 (Van Meijgaard *et al*., 2008) regional climate model (RCM), driven by the ECHAM5 (Roeckner *et al*., 2003) and MIROC (K-1 Model Developers, 2004) general circulation models (GCM), respectively, is analyzed for the

Netherlands and north-western Germany. The gridded, observational E-OBS data set (Haylock *et al*., 2008), version 6.0, for the period 1950-2011 is used for bias correction. The precipitation data have been provided on a 0.22° ×0.22° rotated pole grid and a total of 158 grid points falls into the study area. Figure 1 shows the observed mean annual precipitation totals for the considered grid points. The differences in the annual means are relatively small and no clear spatial pattern is visible.

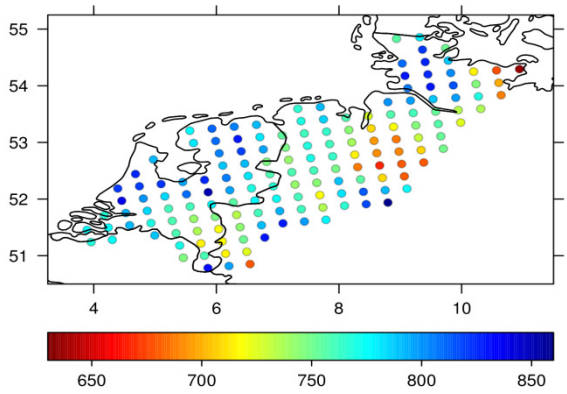


Figure 1: Study area with mean annual precipitation totals in mm

The enhanced greenhouse gas effect is anticipated to be small or not existent in the first decades of the simulation and increasing by the end of the 20th century. This is inconsistent with a simple linear trend over time for the threshold and GPD parameters. Instead of applying more complicated relationships with time, leading to increased estimation uncertainty, a covariate that is considered representative of the enhanced greenhouse gas effect is used. With rising temperatures the water holding capacity of the atmosphere increases. Hence, as extreme precipitation strongly depends on the available precipitable water, compare e.g. Lenderink & Van Meijgaard (2008), temperature is a natural covariate for the non-stationary POT approach. In the following we use the average temperature trend of both climate simulations over the RACMO2 domain as a common temperature covariate.

The selection of the threshold is a crucial step in the application of the POT approach. However, there is still no standard procedure for this, and usually one relies on visual tools. Unfortunately, these rarely give clear indications which quantile should be used for the threshold. We present a threshold selection approach based on the regional setting.

Section 2 outlines the methods. Results and discussion are given in section 3, followed by the conclusion in section 4.

2 METHODS

2.1 Introduction to the peaks-over-threshold model

To study the extremes of independent and identically distributed random variables X_i one can consider the excesses $Y_i = X_i - u$ over a (high) threshold *u*. The Balkema, De Haan, and Pickands theorem states that the distribution of the excesses *Y,* conditioned $Y \geq 0$ oncan be approximated by a generalized Pareto distribution (GPD), if the threshold *u* is sufficiently high and certain regularity conditions hold, see e.g. Reiss and Thomas (2007):

$$P(Y \geq y \mid Y \geq 0) = G_{\xi,\sigma}(y) = 1 - (1 + \xi y / \sigma)^{-1/\xi}, \quad \xi \neq 0, \tag{1}$$

for $y \geq 0$ if $\xi \geq 0$ and $0 \leq y \leq -\sigma/\xi$ if $\xi < 0$ where σ and ξ are the scale and the shape parameter respectively. For $\xi = 0$ the GPD reduces to the exponential distribution. In the case of short-range dependence the GPD approximation applies if one considers the largest value (peak) in a cluster of

exceedances. Several studies considered the GPD also for non-stationary data, using temporally varying parameters, see e.g. Kyselý *et al.* (2010).

2.2 Index flood assumption

Roth *et al.* (2012) introduced an IF approach for non-stationary POT rainfall data, using a time-varying threshold as index rainfall. In this case the IF assumption implies that the dispersion coefficient, i.e. the ratio between the scale parameter and the threshold and the shape parameter are constant over the region of interest but may vary over time, i.e.

$$\sigma_s(t) / u_s(t) = \gamma(t), \quad \xi_s(t) = \xi(t) \tag{2}$$

where $s \in \{1,\ldots,S\}$ and $t \in \{1,\ldots,T\}$ denote the grid point and day. The seasonal mean number λ of the excesses over the threshold in this approach is constant over time and space, which was achieved by using quantile regression to determine the threshold. With (1) we can compute for each site *s* and day *t* the value $r_{s,t}(\alpha)$, that is exceeded on average α times in a season:

$$r_{s,t}(\alpha) = u_s(t)\left(1 - \frac{\gamma(t)}{\xi(t)}\left[1-(\lambda/\alpha)^{\xi(t)}\right]\right), \quad \xi(t) \neq 0 \tag{3}$$

In analogy with a stationary setting, the quantity $r_{s,t}(\alpha)$ is termed the $1/\alpha$-year return level, although $1/\alpha$ gives no longer the expected waiting time between exceedances of $r_{s,t}(\alpha)$.

2.3 Threshold selection

The threshold choice (TC) plot is a widely used graphical tool for the selection of the threshold in the POT analysis. It is based on the fact that once the GPD distribution holds for some threshold it holds for every higher threshold too with the same shape parameter. Therefore, a plot of the estimated shape parameter versus the threshold should be constant. However, owing to the decreasing number of excesses above higher thresholds, the plot eventually becomes unstable and the constant behavior is difficult to see.

For each $\tau \in [0,1)$ we can compute the τ-quantile and the corresponding estimate of the shape parameter, needed for the TC plot. Using the regional setting and averaging both the quantiles and the shape estimates for multiple sites removes some scattering from the individual TC plots. The strength of the spatially averaged plot lies in the increased detection probability of non-constant behavior, when the threshold is too low.

The time-varying threshold can then be determined by quantile regression using the τ selected by the procedure described above. In the following we assume, that the site-specific thresholds may exhibit a common relative trend over the region (for details see Roth *et al*., 2013).

2.4 GPD parameter estimation

For the estimation of the GPD parameters the so-called independence log likelihood, i.e. the log likelihood that would be obtained, if peaks at different sites were independent of each other, is used, see e.g. Moore (1987), Buishand (1991), and Hanel *et al.* (2009). We rely on this simplified likelihood function, because the estimation of the full likelihood function would be virtually impossible, due to the spatial dependence and large dimensionality of the data. This method provides asymptotically unbiased parameter estimates, but the spatial dependence in the data results in a (highly) increased variance of the estimates compared to the variance that would be obtained for independent data. Therefore, Smith (1990) suggested to adjust the standard errors and likelihood ratio tests in a way that is nowadays generalized in the composite likelihood framework, see Varin *et al.* (2011) for an extensive overview. For details of the estimation of the GPD parameters we refer to Roth *et al.* (2012).

2.5 Bias correction

Climate models represent the current status of knowledge about the climate system but are imperfect still. Systematic differences occur between climate model data and observations. For the extremes a correction of these biases can be achieved by adjusting the parameters of the POT model (u, γ, ξ) for the observed biases (for details see Roth *et al.,* 2013). The uncertainty in the projected return levels, due to the unknown GPD parameters, can then be assessed by a bootstrap procedure. Bootstrap samples of the the estimated GPD parameters for the observations and climate model data are generated simultaneously and the estimates for the climate model data are then adjusted for the systematic differences.

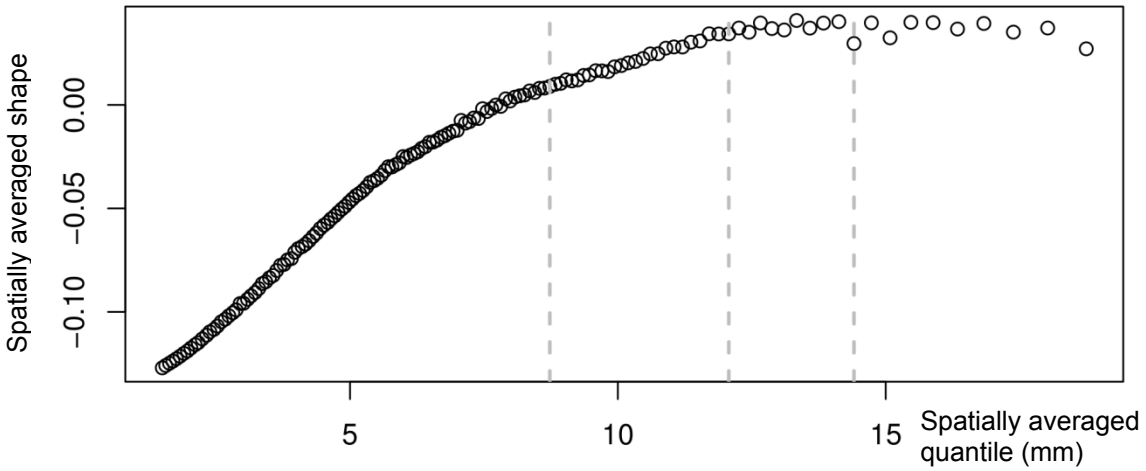


Figure 2: Spatially averaged TC plot for the MIROC driven simulation, winter season. The dashed vertical lines mark the spatially averaged quantile for τ = 0.95, 0.975, and 0.985

3 RESULTS

We illustrate the quantile selection approach proposed in section 2.3, using the winter data from the MIROC driven simulation. From Figure 2 we see that the 95% quantile is too low for the GPD model to hold, i.e. the spatially averaged estimates of the shape parameter to the right of the spatially averaged 95% quantile in the TC plot are not constant. The 97.5% quantile seems to be high enough for the GPD model to hold. Similar pictures for the winter data from the ECHAM5 driven simulation and the observations were obtained. Therefore, we took the 97.5% quantile as threshold for all winter data. For the summer data it was necessary to reject also the 97.5% quantile and the 98.5% quantile was used as threshold.

For winter, the trend in the threshold is significant for both simulations: 2.2% per degree warming for the MIROC driven simulation and 3.6% for that driven by ECHAM5. This results in an increase of roughly 12% and 20%, respectively, by the end of the 21st century. In the summer only the MIROC driven simulation exhibits a significant positive trend of 4.6% per degree warming, corresponding to a 25% increase over the full period. For the precipitation data without a significant trend in the threshold a constant threshold was determined.

Regarding the GPD parameters, it was assumed that the dispersion coefficient varies linearly with the temperature covariate and that the shape parameter is constant. The significance of the trend in the dispersion coefficient was determined by the composite likelihood ratio test, which is an extension of the classical likelihood ratio test that takes the spatial dependence into account (compare Varin *et al*., 2011; Roth *et al*., 2012). In winter, where a significant trend in the threshold was found for both models, the trend in the dispersion coefficient is negligible. In summer the dispersion coefficient is significantly increasing for both simulations. Hanel & Buishand (2011) reported similar trends for the change in a seasonal dispersion coefficient in this region for a larger ensemble of transient regional climate model simulations, based on a regional non-stationary block maxima approach.

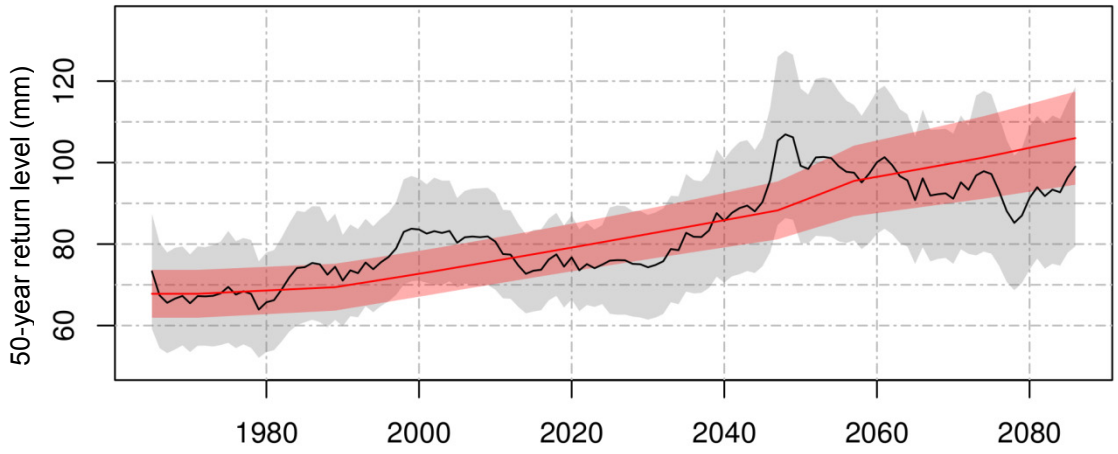


Figure 3: 50-year summer return level, r(0.02), at the grid point closest to De Bilt, for the MIROC driven simulation with 95% confidence bands. The solid black line (respectively gray band) is based on a 30-year moving window and the solid red line (respectively red band) is based on the non-stationary approach

For the selected threshold and the estimated parameters we can compute a time dependent return level using Eq. (3). As an alternative to the non-stationary POT approach, we consider a moving window approach. Therefore, we estimate for each 30-year window of the data (i.e. 1950-1979, 1951-1980, ...) a common shape parameter and common dispersion coefficient, as if the data were stationary. Then the return levels are computed based on these estimates. For the MIROC driven simulation Figure 3 shows, for both approaches, the 50-year return level of daily precipitation in the summer season at the grid point closest to De Bilt, in the center of the Netherlands. The confidence bands are obtained using the asymptotic normality of the maximum independence likelihood estimator of the GPD parameters (Varin *et al.,* 2011). This ignores the uncertainty in the threshold, which is small compared to the uncertainty due to the GPD parameters for the 50-year return level. Overall, the figure shows a good agreement between both methods. However, the 95% confidence band for the non-stationary POT approach is considerably narrower than that for the moving window approach, owing to the increased number of data points used for the estimation. In fact, the relative standard error reduces from about 8.7% for the moving window approach to about 4.5% for the non-stationary approach. Moreover, a monotone trend in the return level is more plausible than the irregular pattern of the trend for the moving window approach, where slightly different selections of the windows can produce quite different estimates of the change, e.g. it matters a lot if the period 2063-2092 (r(0.02) = 85 mm) or the period 2071-2100 (r(0.02) = 99 mm) is taken as future period. This is even more delicate because the control period can be chosen in different ways too.

The 50-year return level of the 1-day summer maximum precipitation near De Bilt from the MIROC driven simulation in Figure 3 is significantly larger than the estimate of 52 mm from the observations. Therefore, a bias correction is needed. Figure 4 shows the bias-corrected 50-year return level for the summer season for both simulations, based on the adjusted threshold and GPD parameters. We see, that in summer the MIROC driven simulation projects a much stronger increase (45%) than the ECHAM5 driven simulation (15%), while the situation in winter is the opposite. Then, the MIROC driven simulation projects a 11% increase in extreme precipitation, and the ECHAM5 driven one a 22% increase.

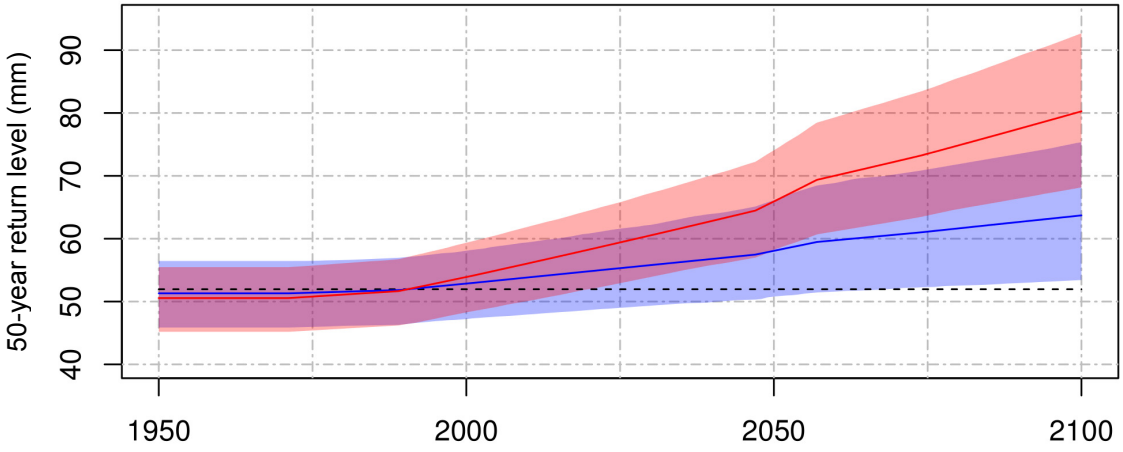


Figure 4: Bias-corrected 50-year summer return level, r(0.02), at the grid point closest to De Bilt, for the MIROC (red) and ECHAM5 (blue) driven simulation with 95% confidence bands. The constant 50-year return level of the observations is given as reference (dashed line)

4 CONCLUSION

In this study we applied the regional non-stationary POT model of Roth *et al.* (2012) to precipitation extremes of two transient climate simulations for the period 1950-2100, conducted with the regional climate model RACMO2, driven by the general circulation models ECHAM5 and MIROC respectively.

Visual inspection of the spatially averaged TC plot leads to a rejection of the 95% quantile as threshold, which is often used in the literature for precipitation data. The non-stationary approach leads to return levels, that are consistent with those obtained by a 30-year moving window approach, but exhibit less uncertainty. For the considered 50-year return level the uncertainty is reduced by a factor of two, compared to the moving window approach. The simulated return levels exhibit a considerable positive bias, which was corrected for by adjusting the parameters of the peaks-over-threshold model. The uncertainty in the adjusted return levels is then strongly governed by the variance of the estimated GPD parameters from the observations. Therefore, accurate estimates of the GPD parameters, based on high quality observed precipitation records, are needed to project future extremes.

Both models project a significant positive trend in the threshold in winter, but not in the GPD parameters. In summer the ECHAM5 driven simulation projects a 15% increase of the 50-year return level, based on a significant increase of the dispersion coefficient. MIROC projects also a significant increase of the threshold, resulting in a 45% increase of the 50-year return level.

5 ACKNOWLEDGMENTS

We thank G. Jongbloed, A.M.G. Klein Tank, and H. van Zanten very much for their help while conducting this work. The research was supported by the Dutch research program Knowledge for Climate. The RACMO2 simulation driven by ECHAM5 and the E-OBS data set were partially funded by the EU FP6 Integrated Project ENSEMBLES (Contract number 505539). The MIROC driven RACMO2 simulation was kindly made available by G. Lenderink and E. van Meijgaard. All calculations were performed using the R environment (http://www.r-project.org).

6 REFERENCES

Buishand, T.A. (1991): *Extreme rainfall estimation by combining data from several sites*. Hydrolog. Sci. J., 36 (4), pp.345-365.

Coles, S. (2001): *An Introduction to Statistical Modeling of Extreme Values*. Springer, London.

El Adlouni, S., Ouarda, T.B.M.J., Zhang, X., Roy, R. and Bobée, B. (2007): *Generalized maximum likelihood estimators for the nonstationary generalized extreme value model*. Water Resour. Res. 43 (3), W03410.

Hanel, M. and Buishand, T.A. (2011): *Analysis of precipitation extremes in an ensemble of transient regional climate model simulations for the Rhine basin*. Clim. Dynam. 36, pp.1135-1153.

Hanel, M., Buishand, T.A. and Ferro, C.A.T. (2009): *A nonstationary index flood model for precipitation extremes in transient regional climate model simulations*. J. Geophys. Res. 114, D15107.

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D. and New, M. (2008): A *European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006*. J. Geophys. Res. 113, D20119.

Hosking, J.R.M. and Wallis, J.R. (1997): *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press, Cambridge, UK.

K-1 Model Developers (2004): K-1 coupled GCM (MIROC) description, edited by Hasoumi, H., and Emori, S., *K-1 Technical Report No. 1*, Cent. for Clim. Syst. Res., Univ. of Tokyo, Tokyo.

Kyselý, J., Picek, J. and Beranová, R. (2010): *Estimating extremes in climate change simulations using the peaks-over-threshold method with a non-stationary threshold*. Global Planet. Change, 72 (1-2), pp.55-68.

Lenderink, G. and Van Meijgaard, E. (2008): *Increase in hourly precipitation extremes beyond expectation from temperature changes*. Nature Geosci. 1, pp.511-514.

Moore, R.J. (1987): *Combined regional flood frequency analysis and regression on catchment characteristics by maximum likelihood estimation*, in Singh, V.P., "Regional Flood Frequency Analysis", Reidel, Dordrecht, pp.119-131.

Reiss, R.D. and Thomas, M. (2007): *Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhäuser, Basel, 3rd edition.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hageman, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U. and Tompkins, A. (2003): *The atmospheric general circulation model ECHAM 5. Part I: Model description.* Technical Report No. 349, Max-Planck-Institute for Meteorology, Hamburg.

Roth, M., Buishand, T.A., Jongbloed, G., Klein Tank, A.M.G. and Van Zanten, H. (2013): *Projections of precipitation extremes based on a regional, non-stationary peaks-over-threshold approach: A case study for the Netherlands and north-western Germany*. Submitted to Weather and Climate Extremes.

Roth, M., Buishand, T.A., Jongloed, G., Klein Tank, A.M.G. and Van Zanten, H. (2012): *A regional peaks-over-threshold model in a nonstationary climate*. Water Resour. Res. 48, W11533.

Smith, R.L. (1990): *Regional estimation from spatially dependent data*. Unpublished manuscript. University of Chapel Hill.

Van Meijgaard, E., Van Ulft, L.H., Van de Berg, W.J., Bosveld, F.C., Van den Hurk, B.J.J.M., Lenderink, G. and Siebesma, A.P. (2008): *The KNMI regional atmospheric climate model RACMO version 2.1.* Technical Report No. 302, Royal Netherlands Meteorological Institute, De Bilt.

Varin, C., Reid, N. and Firth, D. (2011): *An overview of composite likelihood methods*. Stat. Sinica, 21, 5-42.

Near future changes of temperature and precipitation extremes on the regional scale

Katrin Sedlmeier[1], Hendrik Feldmann[1] and Gerd Schädler[1]

[1]Institute for Meteorology and Climate Research (IMK-TRO), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, Email: katrin.sedlmeier@kit.edu

Abstract

The effective drought index and an analogous measure, the effective heat index are applied to two regional climate model ensembles at different resolutions to deduce the near future changes of precipitation and temperature extremes as well as for combined extremes of both variables for an area covering parts of Germany and its surroundings. The ensemble spread is used to derive the uncertainty of the deduced changes.

1 INTRODUCTION

Reliable knowledge about changes of near future extremes is important for impact studies and adaptation and mitigation strategies. Especially compound extremes (two or more extremes occurring simultaneously or consecutively) can have a great impact on society, but have so far received little attention (IPCC, 2012). Furthermore, the associated uncertainties of the change signal as a measure for the robustness of the data are of great interest for decision makers.
For impact studies, the regional trends of the climate change signals are important. Especially in regions with complex topography, climate variables show a high spatial variability and high resolution data is needed (see e.g. KLIWA, 2009; Zolina, 2008; Feldmann *et. al.*, 2012).
Extreme events and especially compound extremes are rare. Using an ensemble of climate simulations the statistical data basis is broader, internal climate variability and modeling errors are sampled and the uncertainties can be quantified by the consistency of the ensemble.
In this study we used a small high resolution ensemble (7 km) and a larger coarser resolved ensemble (25 km) of regional climate simulations to assess the changes of temperature and precipitation related extreme events. The ensemble spread is used to deduce the uncertainty of the findings.
In section 2, an overview of the model and observational data is given before the methods used for the description of extreme events and some aspects of ensemble techniques are introduced in section 3. Section 4 contains some results followed by a discussion in section 5.

2 DATA

2.1 Observational Data

For validation of our model results we mainly used the EObs v.6.0 gridded data set at a resolution of 25km (Haylock *et al.,* 2008). The data set covers land-points in Europe. Temperature and precipitation data are available on a daily basis.
For validation of the high resolution ensemble, the EObs data was bilinearly remapped to the respective grid before the statistics were calculated. For the temperature data, a height correction was applied by multiplying the elevation difference of the remapped EObs elevation and the 7 km grid with a constant lapse rate of 6.5 K/km for each grid point.

2.2 Model Data and Model Region

For impact studies, especially in regions with complex terrain, high resolution data is needed to assess the changes on a regional scale. However, the high computational effort needed for high resolution climate simulations limits the size of the ensemble.

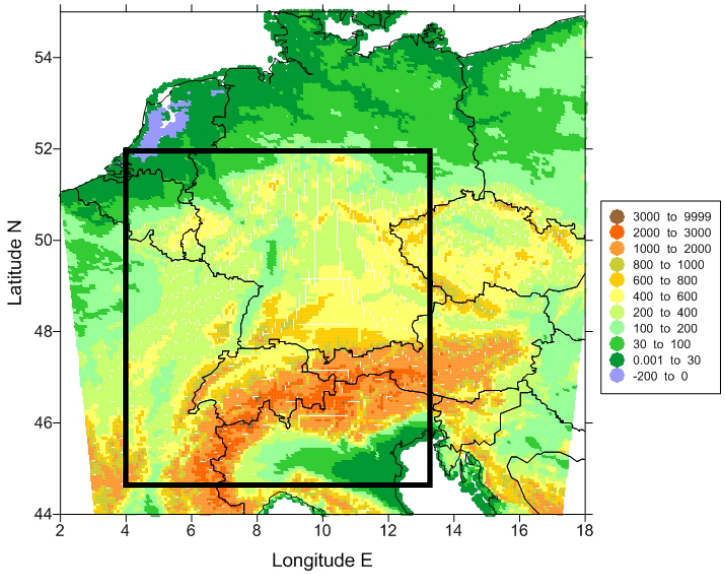


Figure 1: Topography (height [m]) of the CCLM model domain. The black box indicates the investigation area

The regional climate model (RCM) we used for generating our own high resolution ensemble is the non-hydrostatic model COSMO-CLM (COnsortium for Small scale MOdelling model - in CLimate Mode, here abbreviated CLM) which is the climate version of the numerical weather prediction model COSMO of the German Weather service (Doms & Schättler, 2002). Since the difference between the resolution of the driving GCMs and the target resolution of ~7 km is large, a double nesting procedure was used with a coarse nest at ~50 km resolution (covering Europe) and a fine nest at about 7 km resolution covering Germany and Central Europe. The model domain is shown in Figure 1.

Our own CCLM-ensemble (further referred to as E7km), consists of 5 members driven by 3 different global driving models, among them one with different realizations, see Table 1. Since the size of our ensemble is rather small, an enlarged ensemble, using data from the ENSEMBLES project (van der Linden and Mitchell, 2009; http:www.ensembles-eu.org), was analyzed as well. This larger ensemble (further referred to as E25km) additionally includes model runs with 8 different RCM driven by 6 different GCMs adding up to a total number of 21 ensemble members of which all were bilinearly interpolated to the EObs grid before the statistics were calculated for a better comparison of the data. For data interpolated to a finer resolution (RACMO), the temperature data was height corrected. An overview of the data used for the ensembles is given in Table 1. The time periods considered are 1971-2000 (reference period) and 2011-2040 (near future). The investigation area (black box in Figure 1) is smaller than the model area because one member of the E7km ensemble is only available for this area.

Table 1: Model Data for the two ensembles E7km and E25km. Our own data is marked by *, all other data is from the ENSEMBLES project (http://www.ensembles-eu.org/)

RCM\GCM	ECHAM5	Arpege	MIROC	BCM	HadCM3	CC3	CGCM3
E7km							
CCLM	3*				1*	1*	
E25km							
CCLM	3*				1*	1*	
RCA3	2			1	1		
ALADIN		1					
RACMO	3		1				
HIRHAM		1		1	1		
RegCM	1						
VMGO					1		
REMO	1						
CRCM							1

3 METHODS

3.1 Effective drought/heat/hot&dry index (EDI/EHI/EHDI)

For the analysis of extreme events in this paper, the effective drought index (EDI) proposed by Byun & Wilhite (1999) and a comparable measure which we deduced for temperature and call the effective heat index (EHI), are used. A combination of both indices is applied to assess combined hot& dry extreme events. The indices are a measure for the variability of the climate system, describing extremes as deviations from the climatological mean state. Extremes by definition of EDI/EHI are therefore not necessarily "real" extremes in the sense of record breaking events; it is their deviation from the mean that is extreme.
EDI and EHI are calculated for each day d by using equations 1 and 2, they are the standardized anomalies of effective precipitation (further referred to as EP) and effective temperature (ET), respectively.

$$EYI_d = \frac{EX_d - \overline{EX_{d,rm,RP}}}{\sigma(EX - \overline{EX})_{RP}} \quad (1)$$

EP and ET are calculated by a weighted summation over the preceding ds days (see Equation 2). For EP, the same value as in Byun and Wilhite (1999) is used (ds=365, respectively 360 for model runs with only 360 days per year). For the effective temperature, ds was determined as the lag where the autocorrelation function equals 0.5. This was calculated for every grid point for all members of the E7km ensemble and then averaged, leading to a value of ds=47. By using EP/ET the memory effects of soil and atmosphere are taken into account which is especially important when considering droughts.

$$EX_d = \sum_{n=1}^{ds} \left(\frac{\sum_{m=1}^{n} X_{d-m}}{n} \right) \quad (2)$$

In a next step, the climatological mean values of EP/ET are calculated for every day of the year with a running mean over rm days, in the case of $\overline{EP_{d,rm}}$ rm=5 days (as suggested in Byun & Wilhite, (1999)), for $\overline{EP_{d,rm}}$,31 days were used.
By using Equation 1, a time series for the indices is calculated, exemplarily shown in Figure 1 for one member of the E7km ensemble. The climatological mean values of EX and the standard deviation σ were taken from the reference period (marked by RP in Equation 1) for both reference period and near future.

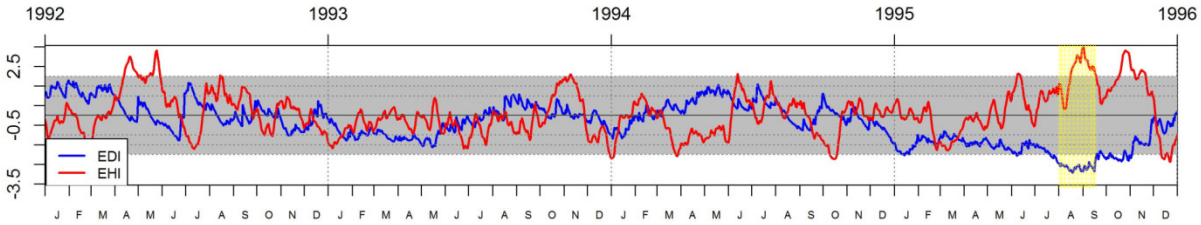


Figure 2: Exemplary EDI (blue) and EHI (red) curves. Curve segments below or above the grey shading have a deviation from the mean greater equal 2σ and are classified as strong extremes. The yellow area marks days where combined extremes occur

The classification of extremes stems from the definition of the indices; they describe the deviation from the climatological mean as a multiple of the standard deviation σ. A moderate extreme is defined as having a deviation greater than σ, a strong extreme greater than 2σ. In Figure 2, the grey shading is a guide to the eye, for points outside of this area EDI/EHI exceed 2σ and are considered as extreme dry/wet (EDI) or hot/cold (EHI). A combined index, which we call the effective hot&dry index EHDI can be derived by combining the threshold conditions for EDI and EHI (highlighted in yellow in Figure 2 is a combined episode of hot&dry extremes).
EDI/EHI/EHDI bear the great advantage that they are independent of linear bias and symmetric, meaning they can be used for wet/dry respectively hot/cold extremes and all combinations.
From the time series different statistics such as the number of extreme days, the number of extreme episodes, the mean and maximum episode length as well as the number of episodes with different durations can be derived where an episode of length x is defined as the occurrence of x consecutive extreme days.
It has to be noted that although we refer to extreme days as hot, cold, dry or wet days this is always a relative measure since it refers to a day with a high positive or negative deviation from the local climatological mean of EP/ET thus assessing the variability. In this paper we will stick to the terms hot, cold, dry and wet, keeping the previous statement in mind.

3.2 Ensemble Methods

Finding the optimal size and composition of climate ensembles is and ongoing research topic. Generally both are limited by the availability of the data due to high computational effort of climate simulations especially for high resolution. Ideally the ensemble members should be independent and they should be drawn from the same distribution as the observations (e.g. Knutti *et al.*, 2010 and Weigel, 2011). If the ensemble spread is too narrow the ensemble is underdispersive and does not cover the whole uncertainty range, if the distribution is overdispersive, the uncertainty range is overestimated. This should always be taken into account when interpreting the estimated uncertainties.

3.2.1 Validation

In this paper only the statistics of the ensemble are validated against the statistics of the observations. For this, we test whether or not the statistics lie within the ensemble spread, later defined as the mean plus/minus one standard deviation σ. General validation of the model temperature and precipitation data is given in the ENSEMBLES project report (van der Linden & Mitchell, 2009; http:www.ensembles-eu.org) and in Berg *et al.* (2013) for the CCLM data.

3.2.2 Ensemble consistency (EC)

Calculation of the ensemble consistency for a climate change signal allows a quantification of uncertainty of the deduced changes which is important information when using data for planning purposes. In this paper, the consistency of a climate change signal is calculated as proposed by Feldmann *et al.* (2012) by subtracting the number of ensemble members with a negative signal (with an absolute value greater than a certain threshold) from the number of ensemble members with a positive signal (greater than a certain threshold) and normalizing it by the total number of ensemble members. An ensemble consistency of 100% /-100% therefore signifies that all members show a positive/negative change whereas a consistency of 0% implies that the models are discordant or that the change is only very small. In order to distinguish these two cases we also calculated the

percentage of ensemble members with no changes or changes smaller than the threshold used for calculation of EC. We call this percentage EC0.

4 RESULTS

The statistics mentioned in the previous section, their climate change signal and ensemble consistency were calculated from the indices for temperature, precipitation and for the combination of both separately for each individual ensemble member and the ensemble mean of all values was calculated.
In this paper, only the results for extreme values of the indices, namely the ones with EDI ≤ -2 (severe drought) respectively EHI ≥ 2 (severe heat) will be shown since these induce the greatest impacts for society. Furthermore only results for hot and dry periods for the hydrological summer half year are shown. As statistics the number of episodes and the mean episode length were chosen since their changes are clear indications of the climate change.

4.1 Validation

In a first step the results for the reference period (1971-2000) were validated as described in section 3.2. Table 2 gives the percentage of grid points where the EObs statistics lie within the ensemble spread. Only grid points where both ensemble and observations are defined were considered. Results are shown both for the E7km and the E25km ensemble.
The percentage of points which could be validated against the observations, i.e. where the observations lie within the ensemble spread, is smaller for the E7km ensemble. This indicates that the ensemble size is too small resulting in too small an ensemble spread. Therefore the E7km ensemble is probably underdispersive and the ensemble consistencies most likely overconfident. For the E25km ensemble more than 70% of the grid points could be validated for all statistics except the mean length of hot&dry episodes.

Table 2: Percentage of EObs observations within ensemble spread

	EObs - E7km	EObs - E25km
Number of dry episodes	66	72
Mean length of dry episodes	45	75
Number of hot episodes	42	76
Mean length of hot episodes	60	80
Number of hot&dry episodes	73	93
Mean length hot&dry episodes	27	35

4.2 Climate change signal of extremes

In the following, the climate change signal between the summer half years of the reference period (1971-2000) and near future (2011-2040) for the number of hot/dry episodes and the mean episode length are presented.
Figures 3-5 show plots of the results for the E25km ensemble. This coarser resolved ensemble was chosen because validation of the E7km ensemble for the reference period (Section 4.1., Table 2) shows that the ensemble spread is rather small and does not contain the observational data for almost half of the grid points (depending on the statistics). In the plots, the grid points where E25km data could not be validated against EObs data, i.e. where the EObs data does not lie within the ensemble spread for the reference period (see Section 4.1), are shaded in grey. All grid points with consistencies EC or EC0 greater than 50 % are marked with a circle (as threshold for the calculation of EC and EC0 we used a relative change signal of 10 % for all statistics). A summary of the findings for both ensembles for grid points which could be validated against the EObs dataset (according to Section 4.1, Table 2) is presented in Tables 3 and 4.

4.2.1 Hot episodes

The climate change signal for the number and mean length of hot episodes is spatially very homogeneous and in the whole investigation area a shift towards more, longer episodes can be observed (Figure 3). There are some regional differences, while the increase of the number of episodes is smaller in northern Italy than in the rest of the area, the increase in the mean length is larger. For all grid points, the ensemble consistency is above 50 %, e.g. the signal is robust. The tendencies of the change signal are the same for the E25km and the E7km ensembles (Table 3) but the magnitude of the range of changes is higher for the higher resolved E7km ensemble (not shown).

Table 3: Percentage of grid points showing an increase or decrease between reference period (1971-2000) and near future (2011-2040) and consistencies |EC| ≥ ±50% or EC0>50% of the change signal for different statistics

hot days and episodes EHI ≥ 2	**Number of hot episodes**		**Mean length of hot episodes**	
	E7km	E25km	E7km	E25km
increase	100	100	100	100
decrease	0	0	0	0
|EC| ≥ ±50% or EC0>50%	100	100	100	100
Dry days and Episodes EDI ≤ -2	**Number of dry episodes**		**Mean length of dry episodes**	
	E7km	E25km	E7km	E25km
increase	100	100	68	34
decrease	0	0	25	58
|EC| ≥ ±50% or EC0>50%	69	16	39	3

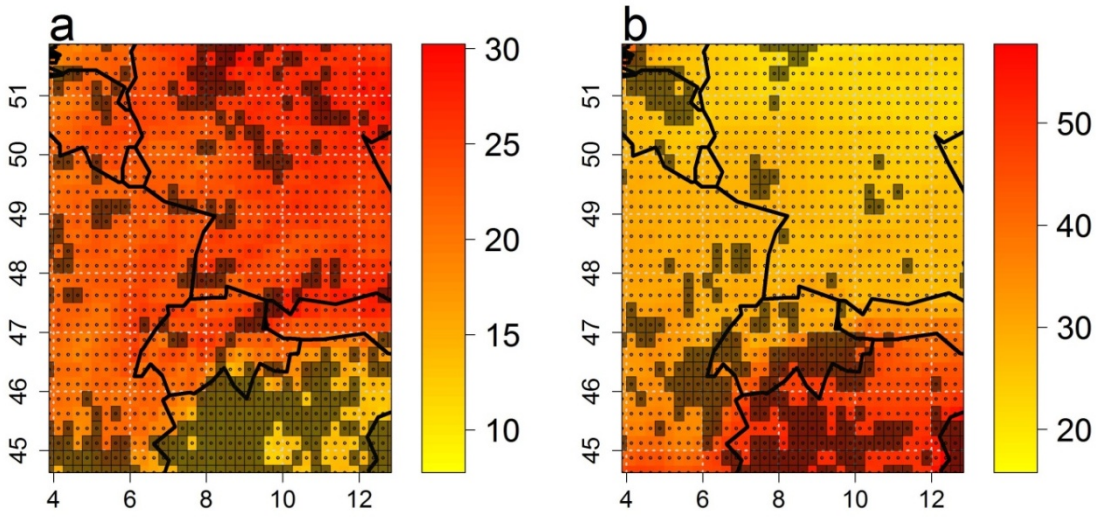


Figure 3: Absolute climate change signal (1971-2000 to 2011-2040, summer half year) for hot extremes (EHI ≥ 2) in the investigation area for the E25km ensemble. a: number of hot episodes, b: mean length of hot episodes. Grid points where the statistics could not be validated against EObs data for the reference period are shaded in grey. Circles mark grid points with |EC| ≥ ±50% or EC0>50%

4.2.2 Dry episodes

As can be seen in Figure 4, the climate change signal for the number and mean length of dry episodes exhibits a higher spatial variability than the temperature signal. In the northwestern part of our investigation area (Central Germany, Belgium and northwestern France) there is mostly a shift

towards more, shorter episodes whereas in the south eastern part (the Alps and northern Italy) most grid points show an increase in both number and mean length of dry episodes. The percentage of grid points where EC>50% or EC0>50% is under 20% for both statistics for the E25km ensemble, especially for the mean length of episodes it is very low. The tendencies of the change signals are the same for the number of episodes for both ensembles, the tendency for the mean length is slightly reversed (Table 3). The magnitude of the range of changes is again higher in the E7km ensemble (not shown). The percentage of grid points with EC>50% or EC0>50% is much higher for the smaller ensemble (Table 3).

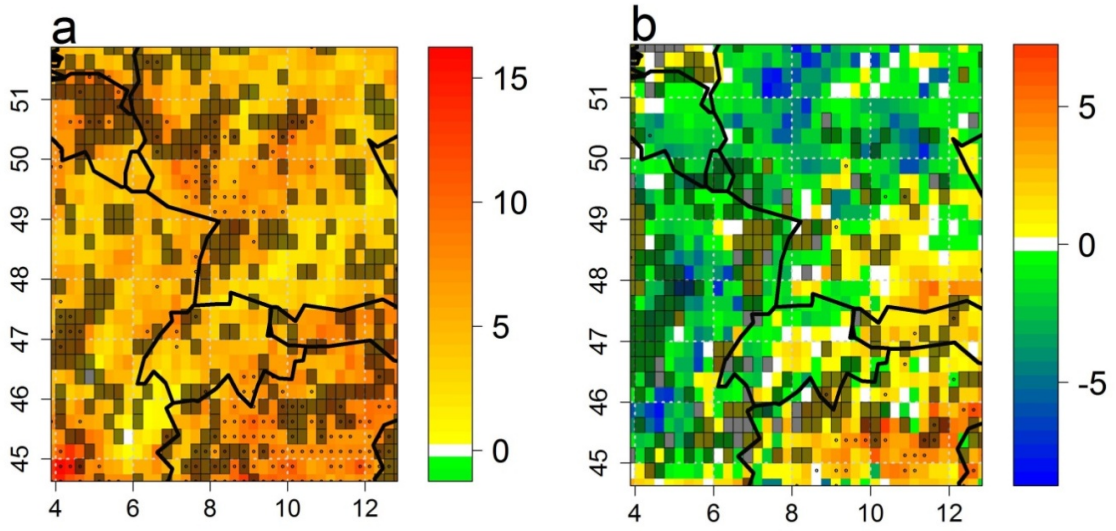


Figure 4: Same as Fig.3, but for dry extremes (EDI ≤ -2)

4.2.3 Combined precipitation and temperature extremes

The results for the change signal of number and mean length of combined hot&dry episodes are shown in Figure 5. The change signal for the number of combined hot&dry episodes is smaller in magnitude than that of the individual statistics. The number of hot&dry days increases in almost all of the investigation area showing the strongest increase in northern Italy and parts of southwestern France. The mean length of the episodes also increases in most of the area. Comparing this to the E7km ensemble (Table 4), the tendencies are very similar but again the magnitude of the range of changes is higher for the higher resolved ensemble (not shown). The ensemble consistencies (EC and EC0) for E25km are high for grid points with a strong change signal, altogether they are under 25% for almost all grid points for both statistics (Table 4). The consistencies for the small E7km ensemble are again much higher.

Table 4: Same as Table 3 but for hot & dry episodes

dry days and episodes EDI ≤ -2 & EHI ≥ 2	**Number of hot dry episodes**		**Mean length of hot dry episodes**	
	E7km	E25km	E7km	E25km
increase	98	97	77	79
decrease	1	0	18	13
|EC| ≥ ±50% or EC0>50%	54	14	43	24

5 DISCUSSION AND CONCLUSIONS

Two ensembles, a 5-member ensemble with 7km resolution and a 21-member ensemble with 25km resolution were analyzed with respect to the climate change signals of hot, dry and hot&dry extremes. These indices describe extremes as deviations from the climatological mean and are thus measures for the changes in variability of the climate system.

Ideally, high resolution data is wanted for impact and planning studies, especially for areas with high spatial variability. However, the validation of our own high resolution E7km ensemble showed that the ensemble size of 5 members is too small. The size of the 21 member ensemble seems to be adequate for describing extremes, in agreement with the findings of Weigel (2008), at least for the extremes of a single variable.
Results for the change signal of the number and mean length of episodes between 1971-2000 and

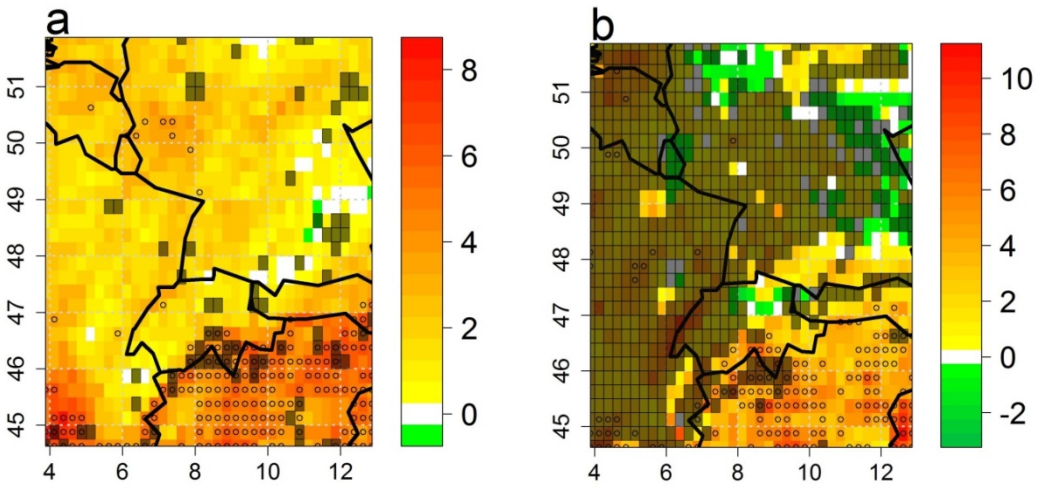


Figure 5: Same as Fig.3, but for hot&dry extremes (EDI ≤ -2 & EDI ≤ -2)

2011-2040 were shown for the summer half year. There is a shift towards more and longer hot episodes in all of the investigation area. The change of dry extremes shows a stronger spatial variability, the number of episodes increases everywhere in the investigation area but the mean length decreases in some parts and increases in others. The number and mean length of combined hot&dry episodes increase in most of the area investigated; only in ~10% the mean length decreases.
We suggest a consistency (EC or EC0) of at least 50% to be necessary for a robust change signal. For the temperature related statistics, all results are robust; the robustness of the statistics for precipitation and combined extremes is only given for part of the grid points for the E25km ensemble. The tendencies of the E7km were confirmed by the larger E25km ensemble. The difference in resolution can be seen by the higher magnitude of the range of the change signals for the E7km ensemble (not shown). At coarser resolution, small scale details are missing and the results are spatially smoother. Especially for precipitation this can play an important role for the results.
We are currently working on enlarging our own high resolution 7km ensemble to obtain an adequate ensemble size and hope to achieve more robust results at higher resolution. Furthermore other statistics, especially for combined extremes will be tested to see whether or not they show similar change signals and consistencies.

6 ACKNOWLEDGEMENTS

This work is funded within the framework of the KLIMOPASS project by the State of Baden-Württemberg, Germany.
The ENSEMBLES data used in this work was funded by the EU FP6 Integrated Project ENSEMBLES (Contract number 505539) whose support is gratefully acknowledged.
We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu)

7 REFERENCES

Berg, P., Wagner,S.,Kunstmann, H. and Schädler, G. (2013): *High resolution regional climate model simulations for Germany:part I—validation*, Clim Dyn, 40 (1/2), pp.401-414.

Byun, H-R. and Wilhite, D.A. (1999): *Objective Quantification of Drought Severity and Duration*, J. Climate, 12 (9), pp.2747-2756.

Doms, G. and Schättler, U. (2002): *A description of the nonhydrostatic regional model LM, Part I: dynamics and numerics,* Consortium for small-scale modelling, Deutscher Wetterdienst, Offenbach, Germany.

Feldmann,H., Schädler, G.,Panitz,H.J. and Kottmeier, C. (2012): *Near future changes of extreme precipitation over complex terrain in Central Europe derived from high resolution RCM ensemble simulations*, Int. J. Climatol. 33 (8), pp.1964–1977.

Haylock, M.R., Hofstra,N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D. and New., M. (2008): *A European daily high-resolution gridded dataset of surface temperature and precipitation*. J. Geophys. Res (Atmospheres), 113 (D20).

IPCC**,** 2012: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change* ,Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582p.

KLIWA (2006): *Langzeitverhalten der Starkniederschläge in Baden-Württemberg und Bayern*, KLIWA-Berichte Heft 8, 94, http://www.kliwa.de/download/KLIWAHeft8.pdf.

Knutti, R., Abramowitz, G., Collins, M., Eyring, V., Gleckler, P.J., Hewitson, B. and Mearns, L. (2010): *Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections*, In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., and Midgley, P.M.," Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections". IPCC Working Group I TechnicalSupport Unit, University of Bern, Bern, Switzerland.

van der Linden and P., Mitchell, J.F.B. (2009): *ENSEMBLES: Climate Change and Its Impacts: Summary of research and results from the ENSEMBLES project*, Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK. 160.

Weigel, A.P. (2011): Ensemble verification. In: Jolliffe, I.T. and Stephenson, D.B.,"Forecast verification: A practitioner's Guide in Atmospheric Science",John Wiley & Sons, Ltd.

Weigel, A.P., Liniger M.A. and Appenzeller, C. (2008): *Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?* Q. J. R. Meteorol. Soc. 134 (630), pp.241–260.

Zolina, O., Simmer, C., Kapala, A, Bachner, S., Gulev, S.K. and Maechel H. (2008): *Seasonally dependent changes of precipitation extremesover Germany since 1950 from a very dense observational network*, J. Geophys. Res. 113 (D06110), 1-17.

Climate-based multivariate Monte Carlo simulation including extremes

Yanira Guanche[1], Roberto Mínguez[1], Fernando J. Méndez[1] and Benjamin P. Gouldby[2]

[1]Environmental Hydraulics Institute IH Cantabria, University of Cantabria, Santander, Spain, Email: guanchey@unican.es

[2]HR Wallingford Ltd., Wallingford, Oxfordshire, UK

Abstract

The design of coastal and offshore structures or the analysis of any wave-driven coastal process requires an accurate wave climate characterization. The availability of long term data series is vital for these purposes. Depending on the location, existing data may be sparse; in these cases, synthetically generated time series offer a practical alternative. Based on a new method to simulate multivariate hourly sea state time series (Guanche et al., 2013b), the main purpose of this paper is to extend this methodology to include the correct simulation of extreme events. This simulation method preserves the statistical characteristics of the existing empirical data. It combines different techniques such as univariate ARMAs, autoregressive logistic regression and K-means clusterization algorithms in order to be able to take into account different time and space scales. The entire process is explained through an example case.

1 INTRODUCTION

Wave climate characterization is required for the assessment of long-term morphological changes along coastlines, designing coastal structures and flood risk assessments. To obtain an accurate characterization of the extreme conditions at a specific location, long-term data sets are required, which are rarely available. Instrumental records have very limited lengths, may have gaps and are not well spatially distributed. Reanalysis data usually provide longer records and avoid missing data and sparse resolution (Reguero *et al.*, 2012; Rascle & Ardhuin, 2013), but the record length is still limited. Synthetically generated time series data may therefore be a useful tool to extend historical data.

To correctly define sea conditions, wave height (H_s) alone is not enough; at least the mean period (T_m) or peak period (T_p) is needed. The dependences between different sea condition variables can be complex. Multivariate models able to capture these dependencies are then of high relevance. Some examples of bivariate models applied to sea conditions can be found in Guedes & Soares Cunha (2000) or Dong *et al.* (2013) and a multivariate approach to characterize sea storm behavior in De Michele *et al.* (2007).

From a different point of view of wind speed scenarios, Morales *et al.* (2010), proposed a multivariate simulation technique based on autoregressive moving average models (ARMA) taking into account the cross-correlations inherent to the problem analyzed.

Data clustering techniques can be used to efficiently manage long time series (Camus *et al.* 2011) and in combination with logistic regression models synoptic patterns of atmospheric or sea conditions can be simulated (Guanche *et al.,* 2013a).

To this end, Guanche *et al.* (2013b) have recently presented a methodology to generate plausible hourly time series of sea state conditions (H_s, T_m and θ_m) by combining clustering techniques with Morales *et al.* (2010) multivariate simulation method and the Guanche *et al.* (2013a) autoregressive logistic regression model. The synthetic time series generated through this method takes into account the atmospheric conditions of the location and the cross-correlation between different variables, but the extreme events simulated are constrained to the historical events.

The simultaneous treatment of point-in-time and extreme-value distributions has been considered by different authors like Coles & Tawn (1991) and Mínguez *et al.* (2012), for example. The aim of this paper is the extension of Guanche *et al.* (2013b) methodology to incorporate parameterized marginal extremes.

2 METHODOLOGY

In this section, the procedure to generate plausible synthetic multivariate sea state time series including extreme events is described. The entire method is divided into three processes: i) the simulation of daily sea level pressure (DSLP) fields, ii) the simulation of daily mean sea conditions (DMSC) and iii) the simulation of hourly sea states (HSS) conditioned to daily mean sea conditions. A general framework of the methodology is shown in Figure 1.

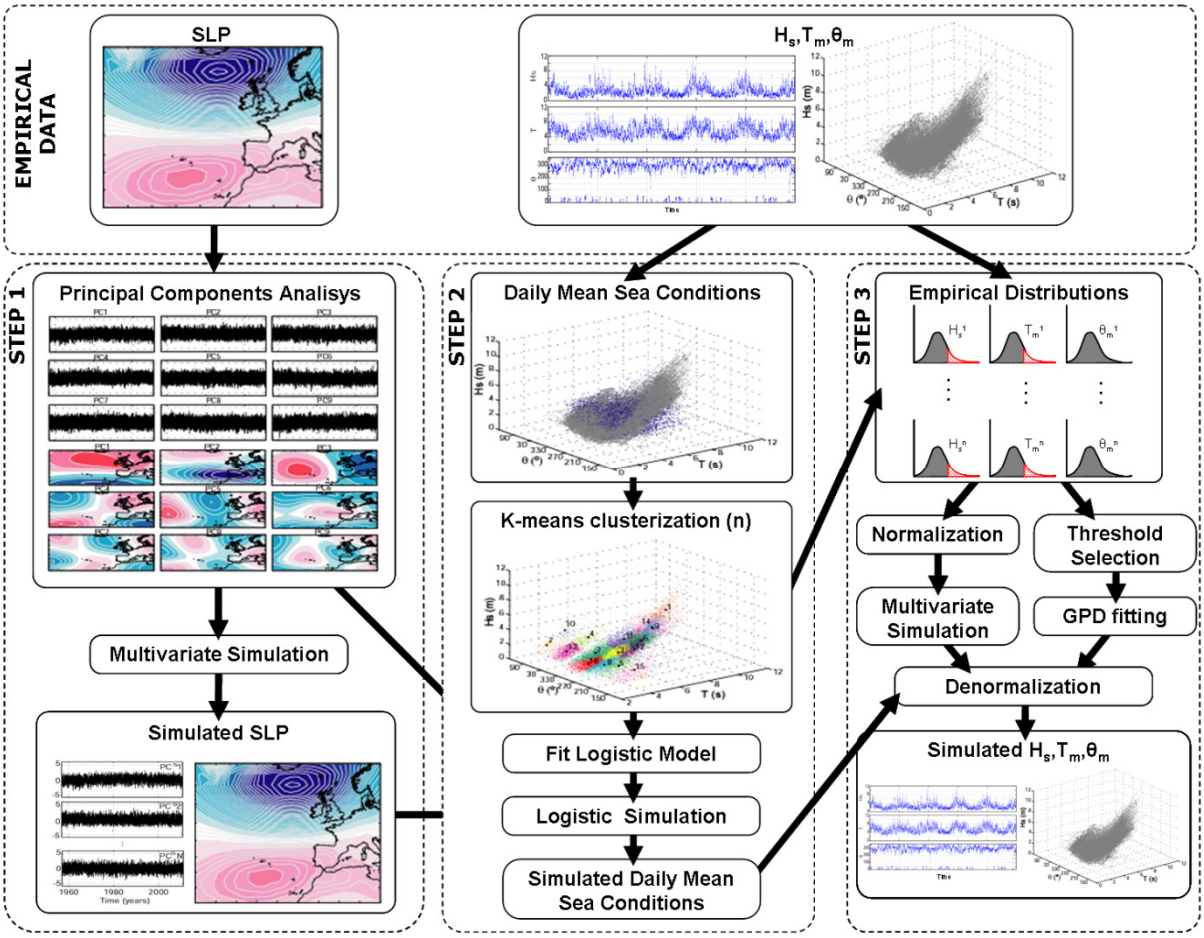


Figure 1: Diagram of the methodology

To implement the method, both SLP and wave time-series (H_s, T_m and θ_m) databases are needed as inputs for the algorithm. SLP fields in the wave generation area of the study location are used as a predictor (covariate) and wave time-series variables are the predictand.

Step 1

Within the first step, the daily averaged SLP fields are simulated. These simulated fields will later be used as explicative variables when simulating daily mean wave conditions. The simulation of SLP fields decomposed into PCs is carried out by using the method proposed by Morales *et al.* (2010), enabling both the autocorrelation of each variable and the cross-correlation between variables.

Step 2

The second step consists of the simulation of daily mean sea conditions. DMSC data is clustered (K-means) into groups, with each group containing data with similar characteristics. Thus, a discrete

time series of daily mean conditions is obtained. The simulation of DMSC uses an autoregressive logistic model (Guanche *et al.*, 2013a). This kind of model enables the consideration of previous states (autoregressive processes) as well as other explicative variables (covariates, in this case SLP PCs).Once the model is fitted using the historical data, the previously simulated DSLP PCs are taken into account for synthetic DMSC time series simulation.

Step 3

Finally, Step 3 of the methodology proposed by Guanche *et al.* (2013b) consists of simulating the variables with an hourly time resolution. Prior to the simulation, the hourly and the daily historical time series are linked, using the original database timestamp, obtaining the empirical distributions of H_s, T_m and θ_m for each cluster defined in Step 2. With these empirical distributions, two analyses are made. Firstly the variables are normalized and then the same multivariate simulation technique (Morales *et al.,* 2010) as in Step 1 is carried on. Then with the empirical distributions, a threshold is selected (Méndez *et al.,* 2006) and GPD distributions are fitted to data above the threshold (Coles & Tawn, 1991; Mínguez *et al.,* 2012). Note that GPD fittings are only done in those clusters with data above the threshold, since only one threshold for each variable is considered. Once the normalized variables are simulated, with consideration of the empirical distributions and the fitted GPD´s, the three variables are transformed back onto the original scale.

Simulated H_s, T_m and θ_m time series present similar marginal and joint distributions as those obtained with the historical data. The disaggregation into daily condition groups improves the simulation results, preserves the seasonality of the variables and allows the consideration of SLP when simulating daily mean sea conditions. Moreover, the inclusion of GPD fits into the right tail of the empirical distribution enables the correct simulation of extreme events.

A more detailed explanation of each step of the process could be found in Guanche *et al.* (2013b).

3 CASE STUDY

3.1 Data

To show the proposed methodology an application on the North West Coast of Spain has been undertaken. Reanalysis databases were used because of their length and consistency. However,, instrumental records may also be used instead where they are available.

SLP data was extracted from the NCEP-NCAR database (Kalnay *et al.*, 1996) covering the area from 25º to 65º N and 52.5º W to 15º E and with a 6-hour temporal resolution from 1957 to 2011. This area covers the wave generation area of the waves arriving to the Northwestern coast of Spain.

Wave data comes from DOW 1.1 (Downscaled Ocean Waves, Camus *et al.,* 2011) from IH Cantabria. This database provides hourly data from 1948-2008 with a spatial resolution of ~200m along the Spanish Coast. The DOW 1.1 database is a downscaled product from GOW 1.1 reanalysis (Reguero *et al*., 2012).

Both data bases share a common period of coverage from 1957-2008. These 52 years of data constitute the input for the method.

3.2 Step 1

SLP fields, daily averaged (DSLP), are decomposed into principal components (PCs), avoiding with this spatially correlated variables and reducing dimensionality. Prior to the PCs analysis, data seasonality is removed by monthly standarizing the data. 14 linearly independent components represent more than 92 % of the variability. These 14 PCs are simulated by using Morales *et al.* (2010) methodology. Figure 2 shows the spatial modes of the 14 PCs.

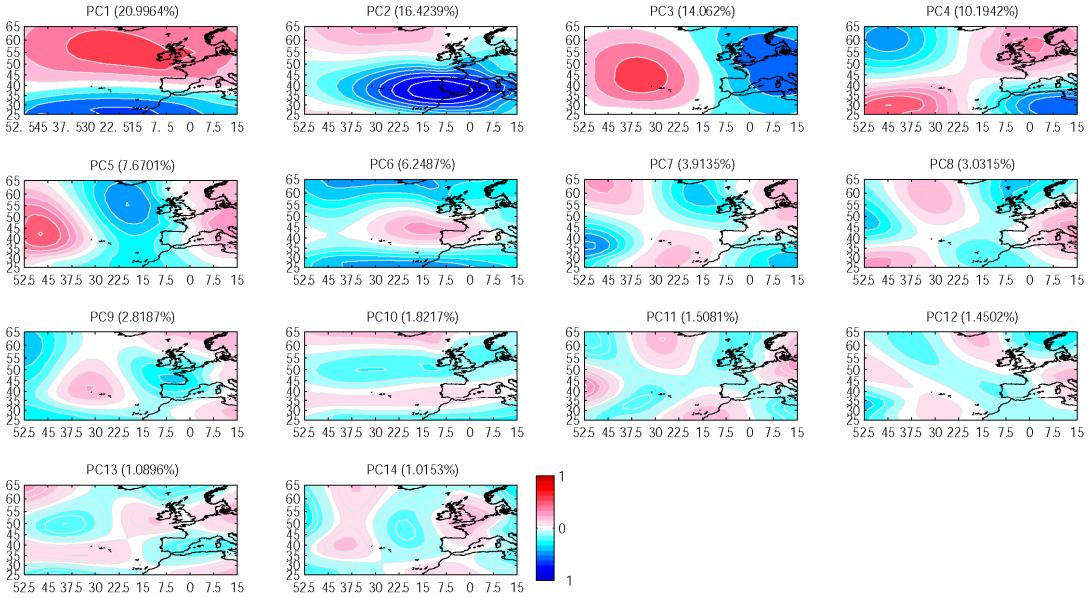


Figure 2: Spatial modes related to the Daily Sea Level Pressure fields Principal Components

Considering the empirical distribution, the 14 PCs are standarized and fitted to a univariate ARMA. In this case the use of an ARMA (2,1) model for all the PCs provided appropriate results. Following the Morales *et al.* (2010) technique, residuals are cross-correlated and the variance-covariance matrix (G) is built. Then, independent standard normal errors are generated and cross-correlated by using G. Finally by introducing these cross-correlated errors within the ARMA models, standarized PCs are simulated. Only lasts the denormalization of the simulated variables. Figure 3 shows the comparison between historical data and simulated data in terms of probability density functions.

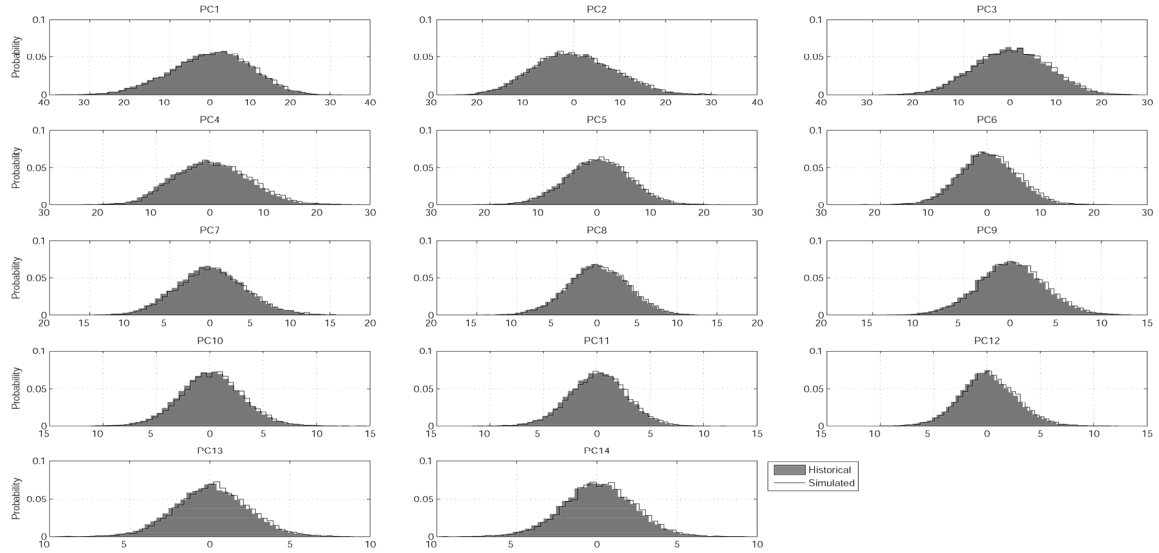


Figure 3: Empirical probability density function related to: i) historical data (grey bars) and ii) simulated data (black line)

3.3 Step 2

Step 2 starts with the clusterization of the trivariate daily mean sea conditions. This is done using a K-means technique. In this case, the data were classified into 16 groups, represented with different colours in Figure 4 and where the centroid position of each group is represented by larger black dots.

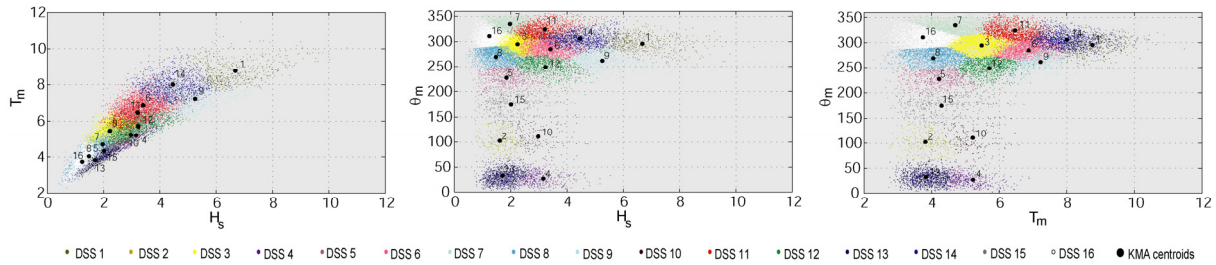


Figure 4: K-means classification of the DMSC

Arbitrarily assigning an integer value between 1 and n=16 for each DMSC in Figure 4, we get the sequence of DMSC, which is the input for the autoregressive logistic model (Guanche *et al.,* 2013a). This model allows the simulation of synthetic sequences of DMSC taking into account different explicative variables, due to its nominative nature. The fitting process of this model is explained in detailed in Guanche *et al.* (2013a) and Guanche *et al.* (2013b). After fitting the model, it takes into account: seasonality, first autoregressive term of daily mean sea conditions, 14PCs of SLP of the concurrent day, 3 first PCs of the day before (first autoregressive term of the DSLP PCs), and 5 first PCs of two days before (second autoregressive term of the DSLP PCs). The fitting process considers the historical DSLP not the simulated obtained in Step 1. Figure 5 shows the comparison between empirical and fitted probabilities for the 16 groups.

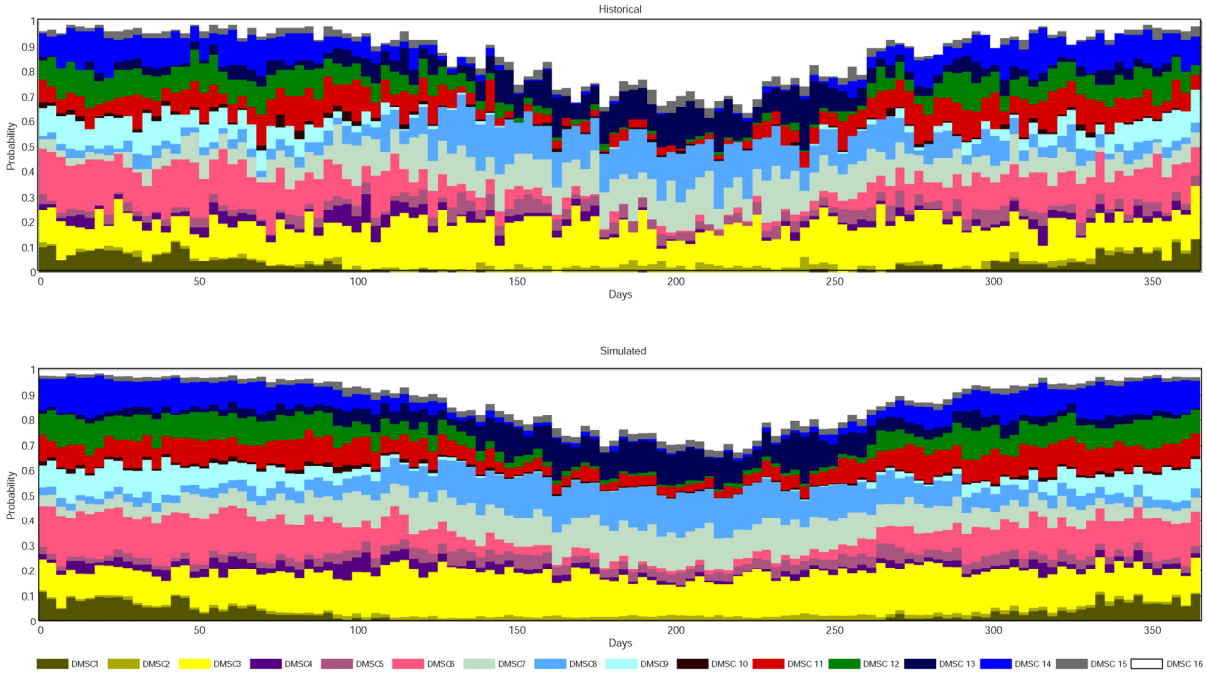


Figure 5: Model fitting diagnostic plot

Once the model is fitted, synthetic sequences of daily mean sea conditions can be generated through a Monte Carlo method, using the simulated PCs as covariates. The ability of the model to reproduce statistically similar sequences of DMSC is checked by developing 100 simulations and evaluating the probability of occurrence of all the groups and the transitions among them. Figure 6a represents the scatter plot of the historical occurrence probabilities of all the groups against the mean and variance of the 100 simulations made. Transitions probabilities among groups are validated in Figure 6b.

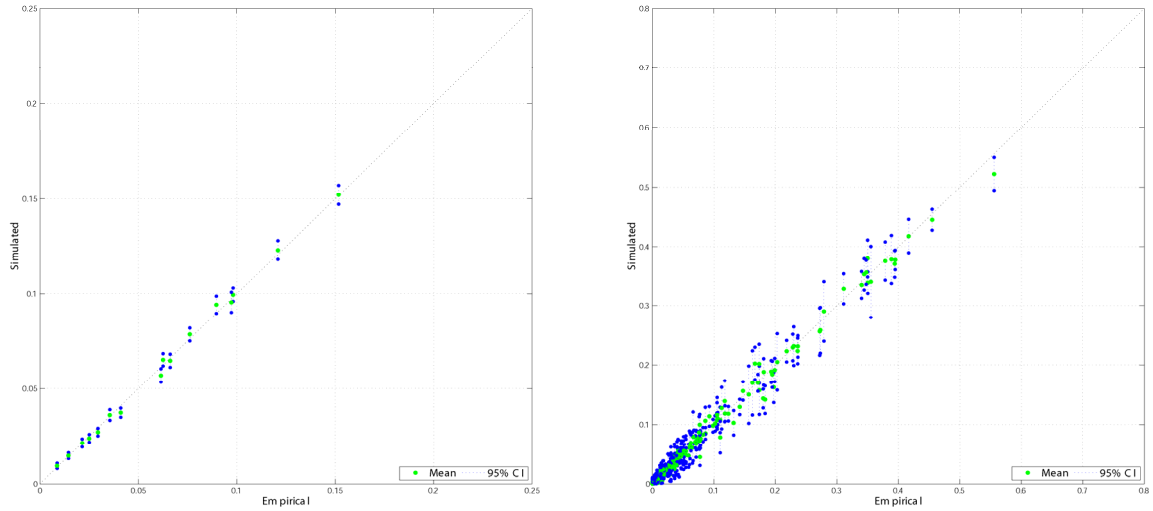


Figure 6: 6a: Occurrence probabilities. 6b: Transition probabilities

3.4 Step 3

Step 3 involves a change in the temporal scale, transferring the daily information into hourly resolution. Initially, the trivariate hourly data are split into the daily mean sea conditions, obtaining n=16 groups for each of the N=3 variables: H_s, T_m and θ_m.

To follow the methodology proposed by Morales *et al.* (2010) and as it was done in Step 1, each variable is normalized. But in this step the normalization is made considering the empirical distribution of the n=16 groups. Then, the 3 normalized variables are fitted to an ARMA model (in this case an ARMA (2,1)) and the residuals are cross-correlated in order to build the variance-covariance matrix G. Once the G matrix is built, normal randomly generated residuals can be cross-correlated and with them the normalized variables could be simulated.

Separately, with the empirical distribution of the three variables, a threshold for two of them (H_s and T_m) is selected. In the case of our study, it is inappropriate to fit an extreme distribution for wave direction because of its circular nature. Following the criteria proposed by Méndez *et al.* (2006), the thresholds were set to be of 9 meters for H_s and 10.5 seconds for T_m. With those thresholds, and imposing the extra condition of having at least 10 data above it, the GPD´s were fitted for all clusters and both variables. Figures 7 and 8 show the empirical distribution of H_s (Figure 7) and T_m (Figure 8) for all the clusters. In those clusters where a Pareto distribution was fitted this adjustment is depicted with a blue line. As seen in Figure 7, from the 16 clusters, only in 2 of them there was enough data above the threshold to fit a Pareto distribution. In the case of Tm, there are 3 clusters with a GPD fitted to its right tail. This occurs because of the method of clusterization chosen; as the Kmeans provides a classification of data with similar characteristics, the extreme events are localized in a few clusters. This allows a better fitting of the right tail of the distributions.

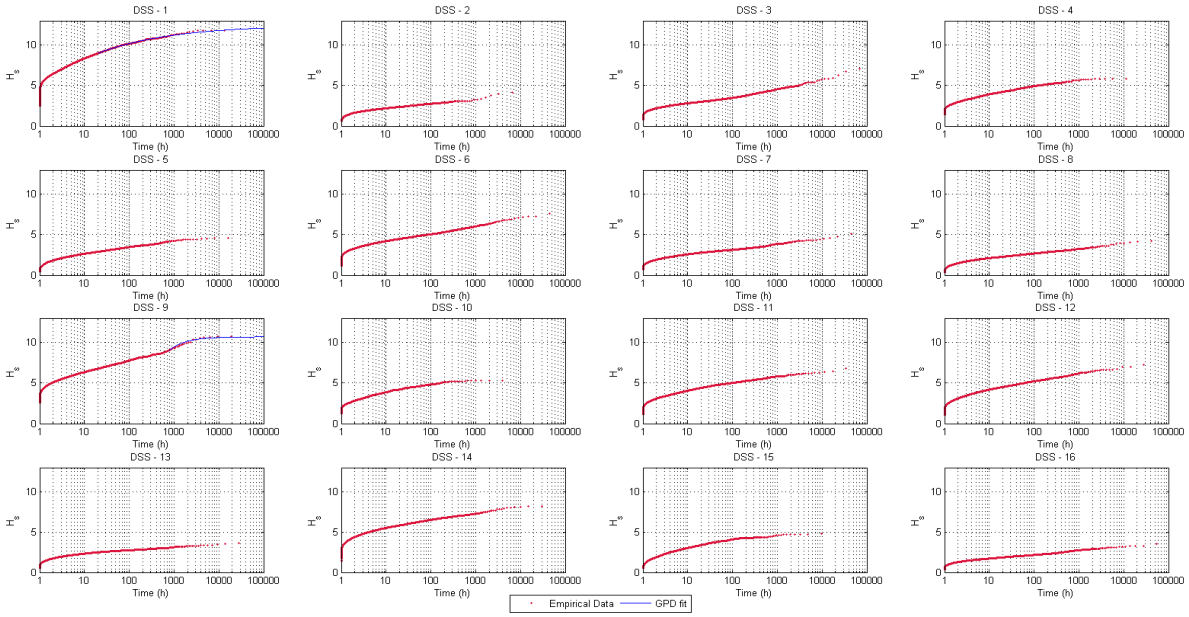


Figure 7: Empirical distribution of H_s for the 16 clusters and GPD fittings

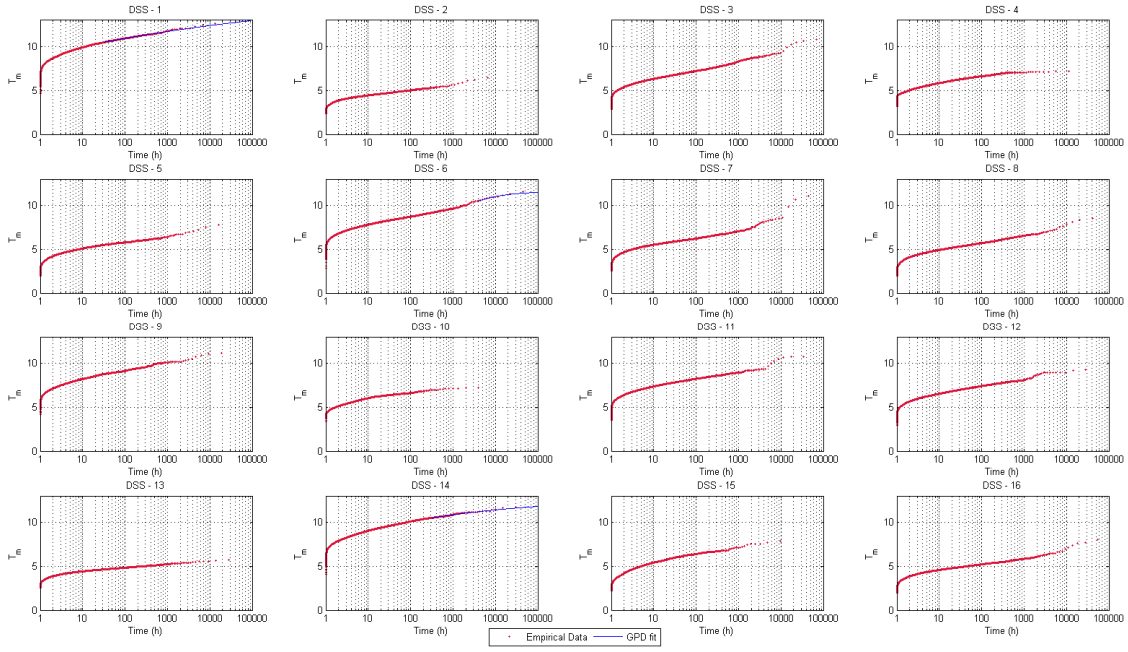


Figure 8: Empirical distribution of T_m for the 16 clusters and GPD fittings

Considering the empirical distributions below the threshold and the GPD fiitings showed in Figures 7 and 8 above them; the denormalization is made (Mínguez *et al.,* 2012). This denormalization is made taking into account the previously simulated sequence of daily mean sea conditions. This way, synthetic hourly trivariate sea states are obtained.

Within Figures 9 and 10 the ability of the proposed methodology to provide good results not only in the joint distributions but also in the marginals is shown. In Figure 9 one simulation of 1000 years is depicted in terms of return period (grey dots) in comparison with the historical record (red dots) for both variables, H_s and T_m. As seen, the marginal distributions are correctly simulated and the inclusion of Pareto distributions to sample extreme events allows the simulation of extreme events different from the historical records but following its same trend. Figure 10 represents the 2-D distributions between the three variables historical vs. simulated. Simulated data maintains the shape of 2-D distributions as well as the joint density distributions (contour lines).

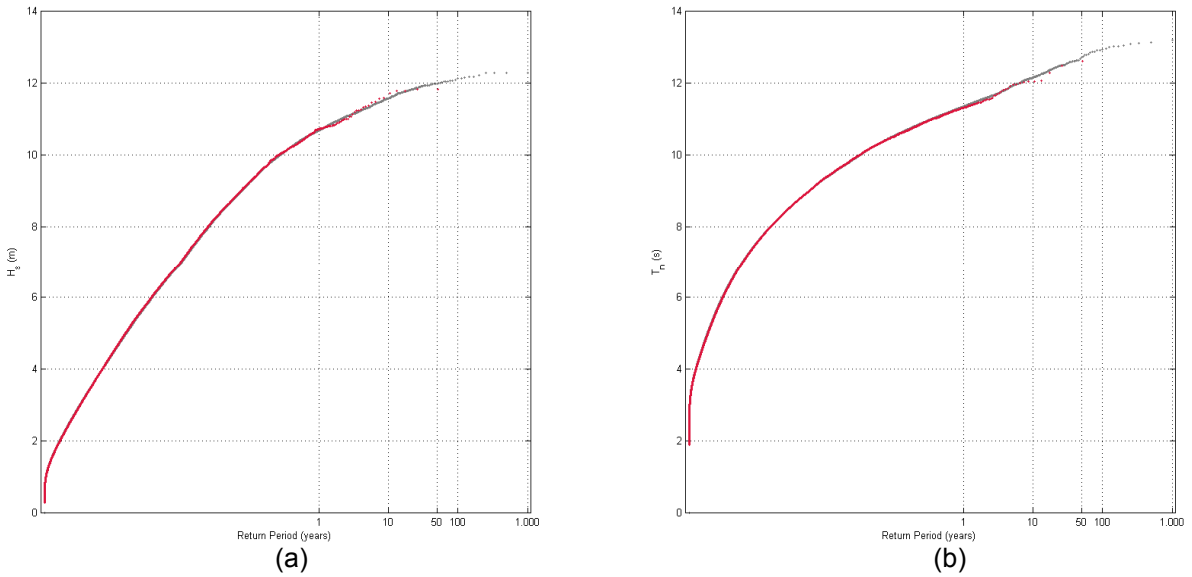


Figure 9: Return period plots of Hs (left subplot) and Tm (right subplot), historical record is plotted in red while a 1000 years simulation is in grey

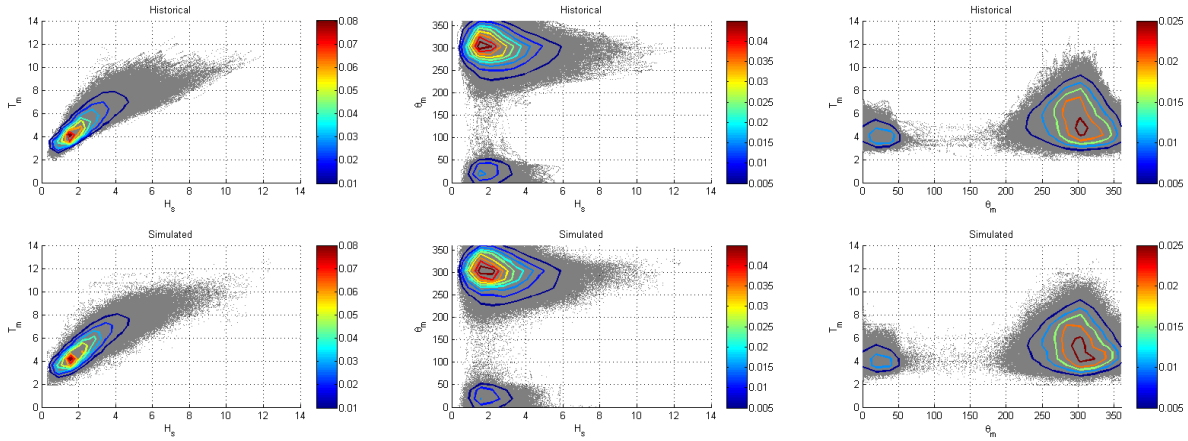


Figure 10: Comparison between historical and simulated joint distributions. Contour lines represent the empirical joint density distribution while dots are hourly data

To check the ability of correctly simulating the seasonality, Figure 11 represents the distribution of the three variables along the year; upper subplots represent the historical seasonality distribution while lower subplots show the results obtained after the simulation process. As seen, the process maintains the seasonal behaviour of the variables.

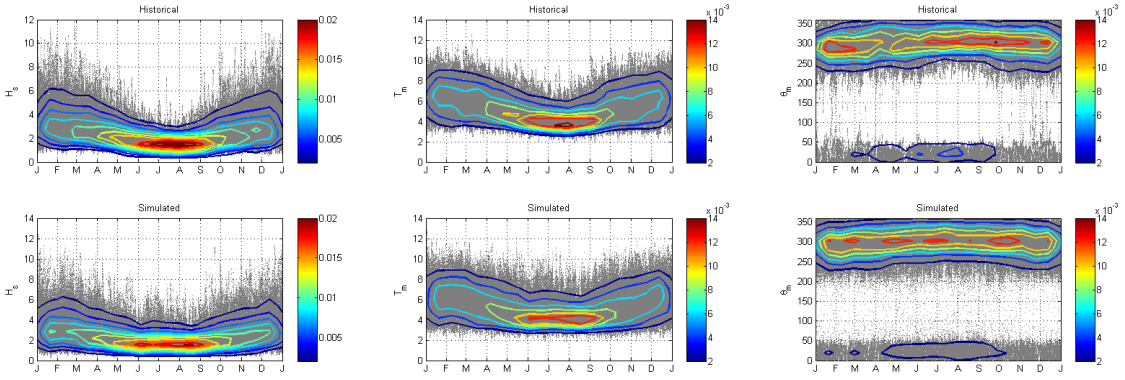
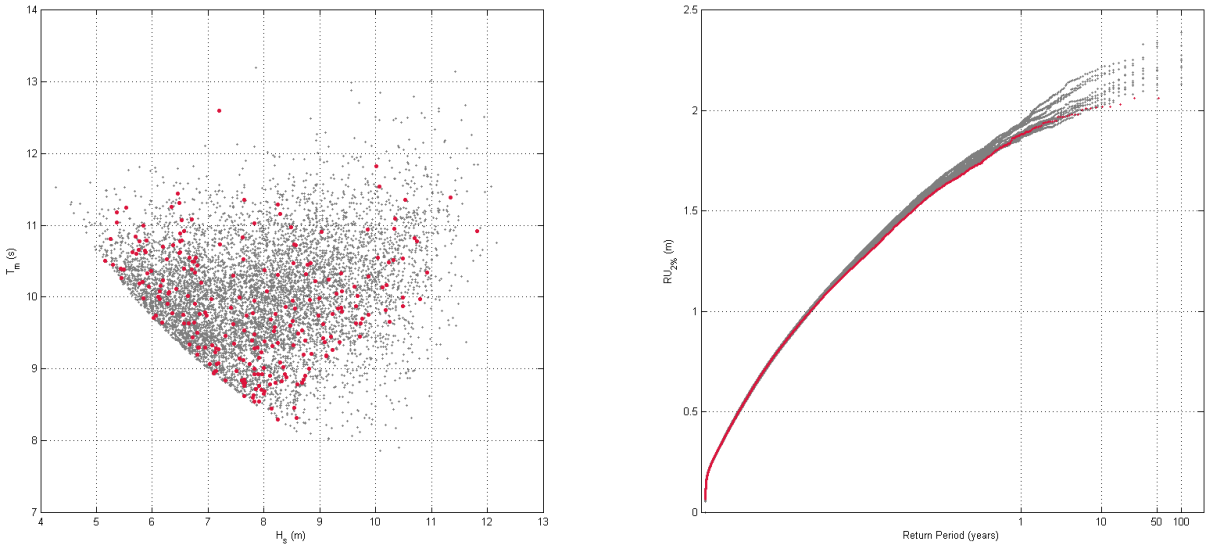


Figure 11: Comparison of the seasonality of historical and simulated. Contour lines represent the empirical joint density distribution while dots are hourly data

Large sample simulation of sea conditions allows an accurate evaluation of extreme events related to those conditions. As an example, considering H_s and T_m it has been calculated the 2% runup on a dissipative beach (Stockdon *et al.,* 2006) for both datasets: i) Empirical data and ii) 1000 years simulation. With the historical distribution of runup, a Peak Over Threshold analysis was performed, setting the threshold at the 99% percentile of the time series (approx. ru=1.27 m) and considering an independency criteria of 3 day separation between events. The extreme events set comprised a total of 235 records, approximately 4.5 events per year. Maintaining the same threshold and independency criteria, extreme events were extracted from the 1000 years simulation time series of runup. H_s and T_m values related to the extracted runup extremes are represented in Figure 12a. Red dots represent the historical extreme events while grey dots represent the simulated extremes. Simulated events cover fairly the space filling the distribution of the limited historical events. In Figure 12b the return period plot of runup is represented; red dots represent the historical distribution while grey dots represent 10 distributions of 100 years lengths. Simulations follow the historical distribution with an increasing variance for higher return periods.



(a) (b)

Figure 12: 12a: Hs and Tm related to extreme events of runup (in red the historical events and in grey the simulated). 12b: Return period plot (historical series in red and simulated in grey)

4 CONCLUSIONS

This work presents an extension of a methodology to reproduce hourly trivariate sea state time series including extreme events. The method combines the use of univariate ARMA models cross-correlated with an autoregressive logistic regression model. The inclusion of extreme events is carried out by fitting the right tail of the empirical distribution to a Generalised Pareto distribution. This combination of techniques allows the simulation of wave climate time series taking into account the different temporal and spatial scales involved.

The possibility of simulating extreme events enhances the utility of the already existing technique. This methodology is therefore applicable for maritime structures design as well as for coastal processes analyses.

5 ACKNOWLEDGMENTS

This work was partially funded by projects “AMVAR” (CTM2010-15009), “IMAR21” (BIA2011-2890) and “PLVMA” (TRA2011-28900) from the Spanish Ministry MICINN. The support of the EU FP7 Theseus “Innovative technologies for safer European coasts in a changing climate”, contract ENV.2009-1, n. 244104, is also gratefully acknowledged. Y. Guanche is indebted to the Spanish Ministry of Science and Innovation for the funding provided in the FPI Program (BES-2009-027228).

R. Mínguez is also indebted to the Spanish Ministry MICINN for the funding provided within the "Ramon y Cajal" program.

6 REFERENCES

Camus, P., Méndez, F.J. and Medina, R. (2011): *A hybrid efficient method to downscale wave climate to coastal areas.* Coastal Engineering 58, pp.851–862.

Coles, S. G. and Tawn, J. A. (1991), *Modelling Extreme Multivariate Events.* Journal of the Royal Statistical Society. Series B (Methodological) 53 (2), pp.377-392.

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

Dong, S., Wang, N., Liu, W. and Guedes Soares, C. (2013): *Bivariate maximum entropy distribution of significant wave height and peak period.* Ocean Engineering 59, pp.86–99.

Guanche, Y, Mínguez. R. and Méndez, F.J. (2013b): *Climate-based Monte Carlo simulation of trivariate sea states.* Coastal Engineering, accepted.

Guanche, Y., Mínguez, R. and Méndez, F.J. (2013a): *Autoregressive logistic regression applied to atmospheric circulation patterns.* Climate Dynamics http://dx.doi.org/10.1007/s00382-013-1690-3.

Guedes Soares, C. and Cunha, C. (2000): *Bivariate autoregressive models for the time series of significant wave height and mean period*. Coastal Engineering 40, pp.297–311.

Kalnay, E.M., Kanamitsu, R., Kistler, W., Collins, D., Deaven, L., Gandin, M., Iredell, S., Saha, G., White, J., Woollen, Y., Zhu, M., Chelliah, W., Ebisuzaki, W., Higgins, J., Janowiak, K.C., Mo, C., Ropelewski, J., Wang, A., Leetmaa, R., Reynolds, R., Jenne, R. and Joseph, D. (1996): *The NCEP/NCAR 40-year reanalysis project*. Bulletin of the American Meteorological Society 77, pp.437–470.

Méndez, F.J., Menéndez, M., Luceño, A. and Losada, I.J. (2006): *Estimation of the long-term variability of extreme significant wave height using a time dependent Peak-Over-Threshold (POT) model.* Journal of Geophysical Research 111, C07024, doi: 10.1029/2005JC003344.

Mínguez, R., Guanche, Y. and Méndez, F.J. (2012): *Point-in-time and extreme-value probability simulation technique for engineering design*. Structural Safety 41, pp.29–36.

Morales, J.M., Mínguez, R. and Conejo, A.J. (2010): *A methodology to generate statistically dependent wind speed scenarios.* Applied Energy 87, pp843–855.

Rascle, N. and Ardhuin, F. (2013): *A global wave parameter database for geophysical applications. part 2: Model validation with improved source term parameterization*. Ocean Modelling http://dx.doi.org/10.1016/j.ocemod.2012.12.001.

Reguero, B.G., Menéndez, M., Méndez, F.J., Mínguez, R. and Losada, I.J. (2012): *A global ocean wave (GOW) calibrated reanalysis from 1948 onwards.* Coastal Engineering 65, pp.38–55.

Stockdon, H. F., Holman, R. A., Howd, P. A. and Sallenger Jr., A. H. (2006): *Empirical parameterization of setup, swash, and runup.* Coastal Engineering 53 (7), pp.573-588.

Increasing Risks for the Management of the North-Eifel Reservoir System caused by Climate Change

Gerd Demny[1], Christof Homann[1], Bernd Hausmann[2] and Matthias Kufeld[2]

[1]Waterboard Eifel-Rur, Düren, Germany, Email: gerd.demny@wver.de

[2]Aachen University, Aachen, Germany

Abstract

The Rur Reservoirs are mainly serving for flood protection, drinking water supply, and water supply for industrial use in the Eifel-Rur region. Within a European INTERREG-IVB-NWE-project studies were conducted, how climate change could affect the reservoir management.

1 THE RUR RESERVOIRS

The river Eifel-Rur is part of the catchment of the international river Meuse. The Rur origins in Belgium, flows through the western part of North Rhine-Westphalia, Germany, and ends in the Meuse in the southern part of the Netherlands (Figure 1). The Rur basin is 2,338 km² (MKULNV, 2005) in area. The construction of the Rur Dam System started over 100 years ago. Today, the waterboard Eifel-Rur (WVER) operates six reservoirs, which are linked with each other. The total volume of the system is 300 Mio. m³ including a volume of 64 Mio. m³ for flood control. The main duties of the Rur Reservoirs are flood protection, provision of raw water for drinking water production and supply of water for industrial use. In addition, the reservoir system serves for water power generation and plays an import role for the regional tourism.

The effectiveness of the dams can be shown on the basis of the following two figures: In case of a flood with a return period of 100 years the discharge in the lower Rur is reduced from 300 m³/s down to 60 m³/s. Under low flow conditions, the reservoirs increase the discharge from 0,33 m³/s up to 5 m³/s. The dams of the Eifel-Rur are also playing an important role for the water resources management of the river Meuse: Although the size of the sub-catchment Rur is only 7 % of the Meuse basin (MKULNV, 2005), the Rur is delivering one quarter of the low flow in the lower Meuse. Therefore the Rur Reservoirs have a significant influence on the drinking water production in the lower Meuse region, especially in addition with the good water quality.

2 THE PROJECT AMICE

In 2009 seventeen institutions from the riparian states of the Meuse have formed the project AMICE (Adaption of the Meuse to the Impacts of Climate Evolutions). The main tasks of the project are the investigations how climate change (CC) affects the water resources of the Meuse and which possible adaption strategies for the water management could be deployed (EPAMA, 2009). The project is coordinated by the French waterboard EPAMA and funded by the European Program INTERREG IVB North-West Europe (NWE).

Figure 2: River basin of Meuse and Eifel-Rur

From the German side the focus is set on the Rur Reservoirs, due to their importance for the regional and superregional water management. The German part of the project is co-funded by the state North Rhine-Westphalia. The main tasks for the studies of the Rur Reservoirs are:

1. How will the climate change in the Rur area in future?
2. Which influence will CC have on the today's management of the Rur Reservoirs?
3. How will the CC change the water related risks at Rur and Meuse (low flow and floods)?
4. Are there any options to adapt the Rur Reservoir management to CC?

The main focus of this paper is set to the general description of these questions. More details can be found in Kufeld *et al.* (2013).

3 QUANTIFICATION OF CLIMATE CHANGE

In the project AMCE two ways of quantification of the CC-effects are carried out: Two periods are selected, one for the near future (2021 to 2050) and one for the far future (2071 to 2100). Both periods are compared with the reference period from 1961 to 1990. Regionalized climate projections were selected for each nation and each region in the Meuse basin respectively. The results of the ZWEK-project (DWD, 2007) were selected for the Rur catchment. Two scenarios for each future period (near and far future) were identified out of the amount of projections in the ZWEK-project: One for representing the driest conditions and one for representing the wettest conditions. These selections are forming the four "national" scenarios. Another set of four scenarios for the near and

far future and for wet and dry climate conditions is produced by area-weighted averaging of all "national" scenarios of each nation in the Meuse basin (Hausmann & Nacken, 2011). This set of four "transnational" scenarios can be seen as the most likely scenarios for the total Meuse basin, but it cannot reflect regional differences.

In Figure 2 the relative change of the annual precipitation is shown for the two different future periods. The wet and the dry scenarios of the national scenarios (blue color) and also of the transnational scenarios (red color) are forming the upper and lower limits for the possible change of precipitation. The maximum increase of annual precipitation is 12.5 % for the far future, the maximum decrease is about -16 % compared to the reference period. In total, the national scenarios for the Rur are showing much wetter conditions than the transnational scenarios for the whole Meuse basin. All scenarios are showing a strong increase of the mean annual air temperature, whether national or transnational. The temperature spans from +2.2 °C to +4.0 °C for the far future.

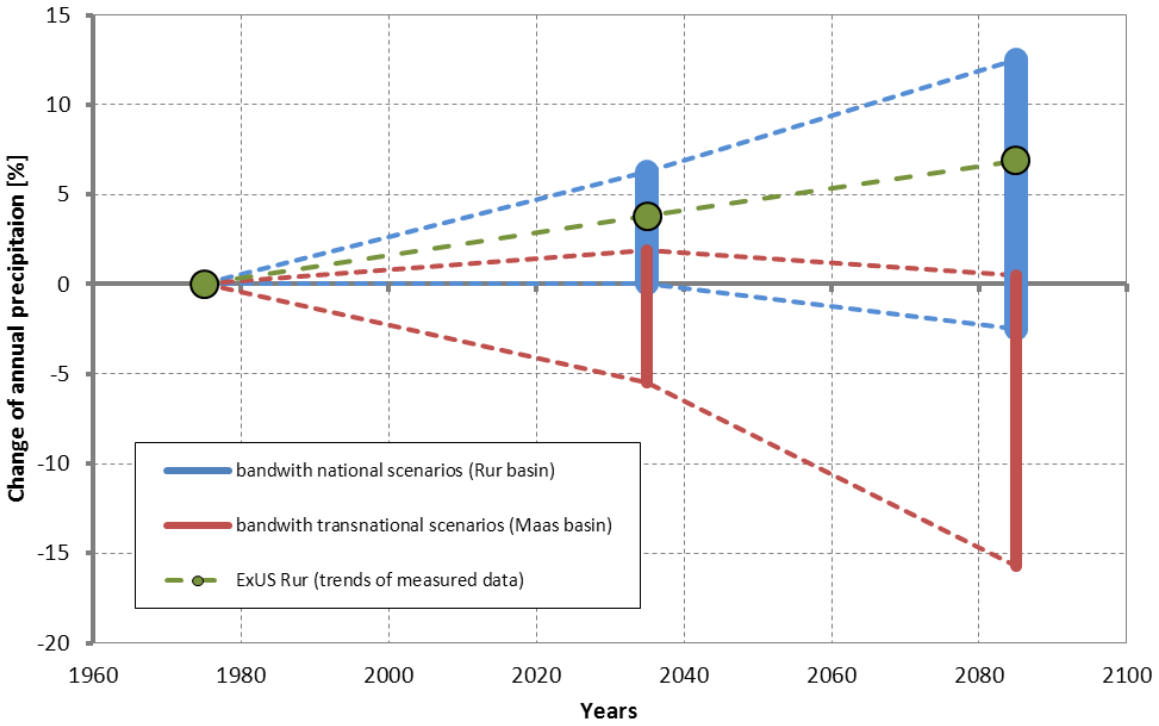


Figure 2: Changes of annual precipitation for the national and transnational scenarios and the ExUS-Rur scenario

The eight scenarios identified are basing on climate model calculations without exception. A very interesting question is, if trends obtained from local measurements in the Rur basin are matching these projections or not. For this purpose data and statistical analyses from the ExUS-project (LANUV, 2010) are complemented with local data of air temperature and precipitation. The additional time series are 58 or 78 years long. The time series of precipitation are showing significant rising tendencies for winter times and weak downward trends for summer times. The total annual precipitation from the observations shows an upward trend. The extrapolation of this trend is shown in Figure 2 (ExUS Rur): The extrapolated measured precipitation is matching to the projections of the national scenarios. So the national Rur scenarios seem to be more suitable for a local CC analysis than the transnational Meuse scenarios. The increase of the annual mean temperature in all national and transnational projection scenarios are confirmed by the trend analysis of the measured data: The time series of temperature in the Rur basin show a strong upward trend to +3.2°C for the far future.

In summary all projections show a change of climate in the Rur basin. The national scenarios could be certified by the analysis of measured time series. The precipitation will increase in winter times and slightly decrease in summer times. The upward trends of temperature could also be confirmed by the trend analysis of measurements. The changes in precipitation and temperature will have an influence on the water balance in the Rur basin and will have an impact on the Rur Reservoir management. The trends identified from the CC calculations and from the measured data show a

broad band width. So the quantification of CC in the regional scale is still afflicted with a high uncertainty.

4 CC AND RESERVOIR MANAGEMENT

The eight national and transnational scenarios from the climate projections together with the two ExUS-Rur scenarios are applied to the Rur Reservoir simulation. The total catchment of the Upper Rur Reservoirs is modeled with five runoff-models (NASIM) and the reservoir management is represented by two management models (TALSIM). The total of seven models is linked together to one model calculation chain. The simulation of the Lower Rur basin below of the reservoirs down to the outlet into the Meuse is calculated by additional three catchment models (NASIM). Origins for the calculations are the daily and hourly time series of precipitation, air temperature and potential evaporation of the reference scenario (period 1961 to 1990). For the CC calculations the delta approach is applied: The time series of the reference scenario are manipulated by the quarterly and monthly differences between each of the eight CC scenarios and the reference period. The advantage of the delta approach is the fact, that no bias, e.g. from the climate model calculation, can have an effect on the results of the reservoir calculation.

Figure 3 shows a detail of the discharge duration curve of the lowest Rur dam. The flow duration curves of the reference scenario and the five scenarios for the far future (2071 to 2100) are represented. For nearly all scenarios the minimum discharge of 5 m³/s into the lower Rur can be guaranteed, the longest undercut lasts two days in the 30 year time period. Only the transnational dry scenario shows a violation of the minimum discharge of 6 days every year. The results for the scenarios for the near future (2021 to 2050) are similar to the results for the far future, but the undercuts are of minor importance here.

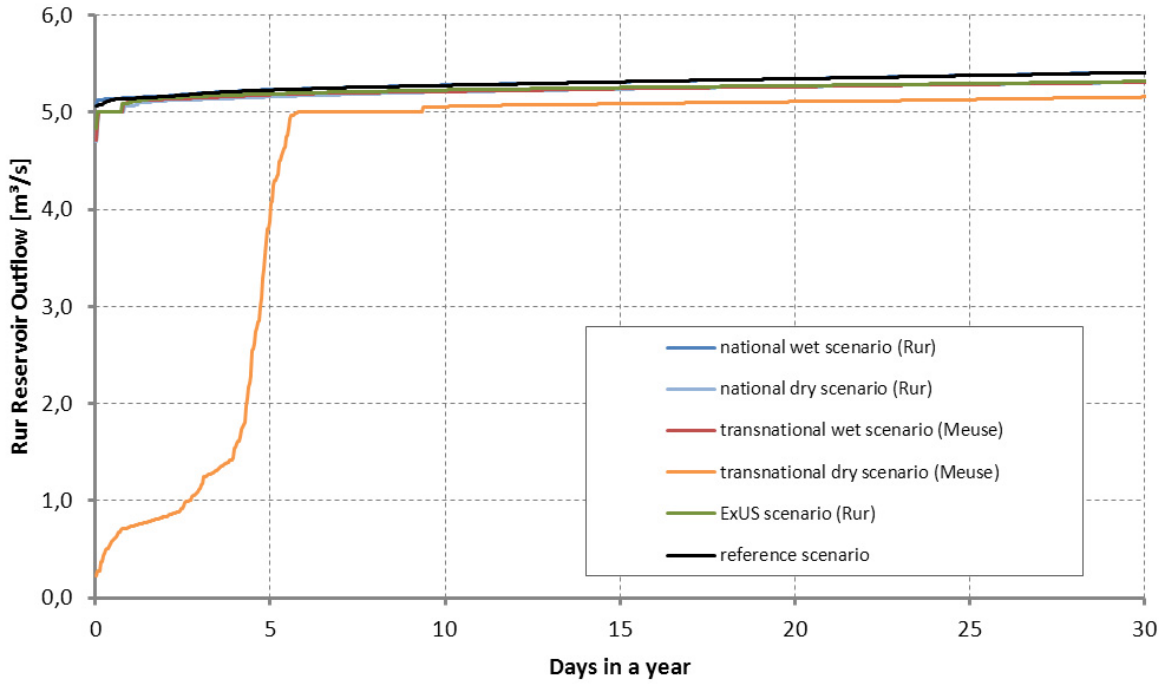


Figure 3: Flow duration curves of the discharge of the Rur Reservoir System for the far future (2071 to 2100)

The flood risks of the Rur Reservoir System are calculated by using the flood characteristics simulation (LANUV, 2004). The characteristics of measured and calculated flood waves are identified by using statistical methods. Together with statistics of the calculated filling levels of each reservoir an amount of 10,000 flood events is generated and applied to the whole Rur Reservoir System. The results are empirical series of the flood discharge into the lower Rur for each CC scenario. The national and the ExUS-Rur scenarios are showing partially a severe overstepping of the limit of 60 m³/s flood discharge into the lower Rur (Figure 4). For both scenarios, the return period of a discharge of 60 m³/s will decrease from 200 years down to 60 and 50 years respectively.

The 100-year flood will increase to 80 m³/s and will affect the flood risks downwards the reservoir system. Also the extraordinary flood event of a return period of 1,000 years will increase up to factor 1.6 in the worst case. The results for the far future are showing the same tendency.

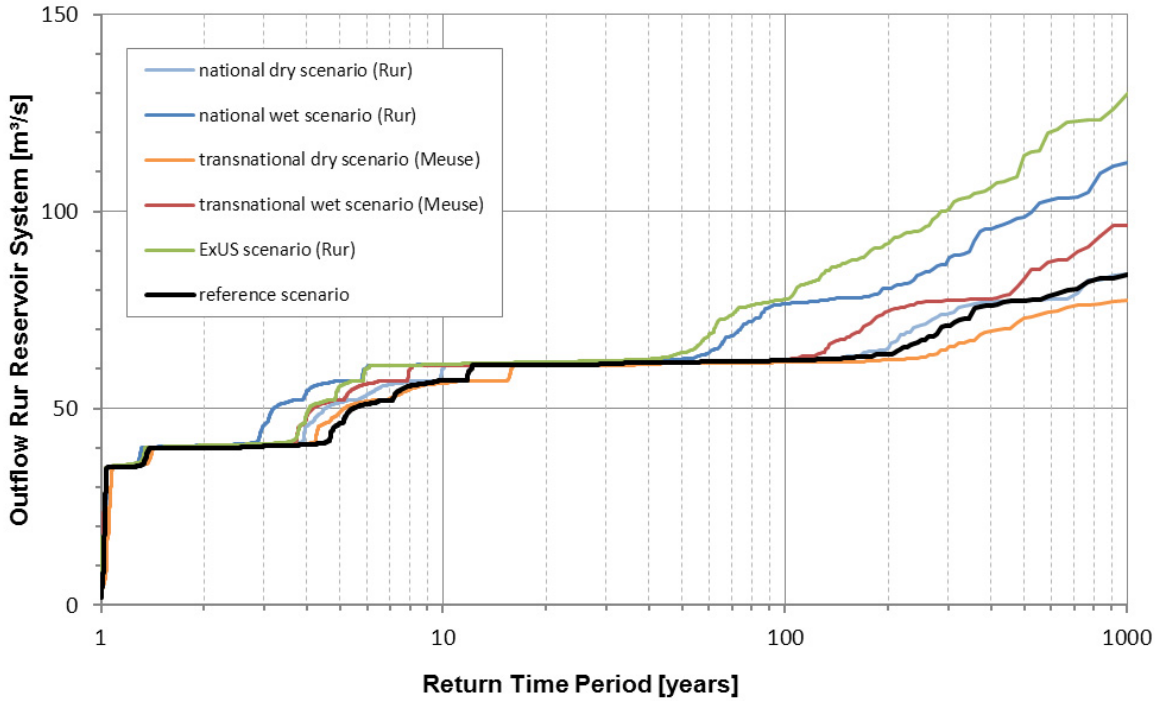


Figure 4: Empirical series of flood discharge from the Rur Reservoirs into the lower Rur for the near future (2021 to 2050)

A set of measures is identified, how to adapt the Rur Reservoirs System to the CC. It can be shown, that especially the increase of the volume for flood control in combination with more restricted rules for the outflow under low flow conditions can help to bridge the gaps due to the CC effects.

5 CONCLUSION

The analysis of different CC scenarios for the Rur basin shows a strong trend to a higher yearly mean air temperature (+2.2 to 4.0°C) up to the year 2100. The trends of precipitation are showing some differences: For the whole Meuse basin, the annually precipitation will drop down to -15.7 % for the worst case, for the Rur basin an increase up to 12.5 % is projected. The analysis of measured time series in the Rur basin confirms the rise in temperature and in precipitation. Within a year, the projected precipitation shows the tendency to increase in winter and to decrease in summer.

A set of ten scenarios is identified from the CC calculations and the measured data for the near and the far future. These scenarios are applied successfully to the runoff and water balance models for the six Rur reservoirs. Under low flow conditions, only one of the ten scenarios show a severe undercut of the required minimum discharge of 5 m³/s. Under flood conditions, several scenarios lead to an exceeding of the maximum allowed discharge (60 m³/s) from the dam system into the lower Rur. In summary, the Rur Reservoir System seems to be more robust against dry climate conditions than against wet conditions.

An analysis of measures shows, that it is possible to adapt the Rur Reservoir System to the CC conditions. The uncertainty of the CC projections is still very high, so it is not advisable to take specific actions by now. But the present study and the developed methods will help to evaluate future changes in the reservoir management and to identify “no regret” measures.

6 ACKNOWLEDGMENTS

This work was funded by the European Program INTERREG IVB North-West Europe (NWE) and co-funded by the state North Rhine-Westphalia, Germany.

7 REFERENCES

DWD (2007): *ZWEK - Zusammenstellung von Wirkmodell-Eingangsdatensätzen für die Klimafolgenabschätzung,* Offenbach, Germany, (http://mud.dkrz.de/projects-at-md/sg-adaptation/other-regional-model-data/zwek/index.html, call 06.03.2013).

EPAMA (2009): AMICE Meuse Maas, France, (http://www.amice-project.eu/de/index.php, call 06.03.2013).

Hausmann, B. and Nacken, H. (2011): *Mögliche Auswirkungen des Klimawandels auf das Abflussregime der Eifel-Rur,* in: Hydrologie & Wasserwirtschaft – von der Theorie zur Praxis, Beiträge zum Tag der Hydrologie am 24./25. März 2011 an der Technischen Universität Wien, Forum für Hydrologie und Wasserbewirtschaftung, Heft 30.

Kufeld, M. et al. (2013): *Anpassung des Talsperrenbetriebs an die Klimaänderung am Beispiel des Rurtalsperrensystems*, WasserWirtschaft, issue 5.

LANUV (2004): *Ermittlung von Bemessungsabflüssen nach DIN 19700 in NRW*, Merkblätter Band 46, Düsseldorf, Germany, (http://www.lanuv.nrw.de/veroeffentlichungen/merkbl/merk46/merk46start.htm, call 06.03.2013).

LANUV (2010): *Extremwertstatistische Untersuchung von Starkniederschlägen in NRW (ExUS), Abschlussbericht,* Düsseldorf, Germany, (http://www.lanuv.nrw.de/klima/pdf/ExUS_Bericht_1a.pdf, call 06.03.2013).

MKULNV (2005): *Ergebnisbericht Rur und südliche sonstige Maaszuflüsse, Wasserrahmen-richtlinie in NRW – Bestandsaufnahme*, Düsseldorf, Germany, (http://daten.flussgebiete.nrw.de/-bestandsaufn/daten/maas_sued/index.html, call 06.03.2013).

Previously published in the fwu-series (Mitteilungen des Forschungsinstitus Wasser und Umwelt der Universität Siegen)

Issue No.	Title	Author(s)	Year
1	Untersuchungen zur Ermittlung von hydrologischen Bemessungsgrößen mit Verfahren der instationären Extremwertstatiskik - Methoden und Anwendungen auf Pegelwasserstände an der Deutschen Nord- und Ostseeküste	Mudersbach, Christoph	2010
2	CoastDoc 2010 - Beiträge zum 1. Doktorandenseminar CoastDoc, Universität Siegen	Jensen, Jürgen (Editor)	2011
3	Expertenseminar Watershed Management and Rural Sanitation	Bormann, Helge (Editor)	2012
4	Statistical methods to assess the hydrodynamic boundary conditions for risk based design approaches in coastal engineering - Methods and application to the German North Sea coastline	Wahl, Thomas	2012
5	Towards sustainable water quality management	Bormann, Helge und Althoff, Ingrid	2013
6	Proceedings of the 1st International Short Conference on Advances in Extreme Value Analysis and Application to Natural Hazards (EVAN 2013)	Jensen, Jürgen (Editor)	2013

Herausgeber:
Forschungsinstitut Wasser und
Umwelt (fwu) der Universität Siegen
Paul-Bonatz-Straße 9-11
57076 Siegen